

Influence of the Conformation of Conotoxins on their Bioactivity

Dissertation
zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Pascal Heimer

aus

Mayen

Bonn, März 2017

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Die vorliegende Arbeit wurde in der Zeit vom Januar 2012 bis Januar 2017 am Institut für Pharmazeutische Chemie der Rheinischen Friedrich-Wilhelms-Universität Bonn unter der Leitung von Prof. Dr. Diana Imhof angefertigt.

Erstgutachter: Prof. Dr. Diana Imhof

Zweitgutachter: PD Dr. Jochen Winter

Tag der Promotion: 08.06.2017

Erscheinungsjahr: 2017

*Für meine Eltern, Familie, Freunde und
Carolin für die jahrelange Unterstützung*

Es könnte alles so einfach sein, ist es aber nicht!

– *Die fantastischen Vier*

Abstract

In the last decade peptides filled the gap between small molecule drugs and biologics as therapeutics. A rich source of such peptides represents the venom of different animals evolved over millions of years. Conotoxins are small, disulfide-rich neuropeptides isolated from cone snails of the genus *Conus*, which act on different biological targets like ion channels or receptors. These toxins contain a complex cocktail of bioactive substances and are utilized for self-defence or hunting prey. With over 700 species known, these marine invertebrates reveal a large potential of novel pharmacologically active molecules, e.g. for the treatment of severe and chronic pain. Although peptide synthesis is established nowadays, preparation and characterization of disulfide-rich peptides is still a challenge. The formation of different disulfide isomers is one main issue facilitating the elucidation of such disulfide-rich peptides and proteins.

The aim of this work was to investigate the folding of small multiple disulfide-bridged peptides under different conditions and perform structure-activity relationship studies of the resulting products. Within the scope of this project different conotoxins were synthesized by solid-phase peptide synthesis. Disulfide formation was achieved by different methods i.e. oxidative self-folding, oxidation in Ionic Liquids and successive oxidation in combination with an orthogonal protecting group strategy. The oxidized conotoxins were systematically investigated and characterized with different analytical methods. Elucidation of the final three-dimensional structure was performed by NMR spectroscopy. Furthermore, determination of the disulfide connectivity was investigated by different methods of mass spectrometric analysis (LC-ESI and MALDI) after partial reduction and derivatization. The results of all applied analytical methods revealed their advantages and limitations for the discrimination of different disulfide isomers within conotoxins depending on primary amino acid sequence. Biological testing of the synthesized peptides on ion channels was performed in parallel with computational studies to enlighten the structure-activity relationships. The work presented in this thesis significantly extends the knowledge about conotoxins and disulfide formation, their impact with respect to biological activity, and the challenging characterization and structure elucidation procedures.

Zusammenfassung

In den letzten Jahrzehnten haben Peptidtherapeutika begonnen, die Lücke zwischen niedermolekularen Arzneimitteln und Biopharmazeutika zu schließen. Eine reiche Quelle für solche Peptide stellen die Giftstoffe verschiedener Tiere dar, die sich über Millionen von Jahren entwickeln konnten. Conotoxine sind kleine, disulfidverbrückte, neuroaktive Peptide, die verschiedene biologische Ziele wie Ionenkanäle und Rezeptoren beeinflussen und aus dem Gift der Kegelschnecke der Gattung *Conus* isoliert werden. Dieses Gift besteht aus einem komplexen Cocktail aus bioaktiven Substanzen und wird von der Schnecke zur Selbstverteidigung oder zur Jagd eingesetzt. Mit über 700 bekannten Spezies bergen diese wirbellosen Meerestiere ein großes Potential für neuartige, pharmakologisch wirksame Moleküle, z.B. zur Behandlung von akuten und chronischen Schmerzen. Obwohl die Peptidsynthese heutzutage etabliert ist, stellt die Herstellung und Charakterisierung von disulfidverbrückten Peptiden immer noch eine Herausforderung dar. Hierbei ist die Bildung von unterschiedlichen disulfidverbrückten Peptidisomeren ein Hauptproblem, welches die Aufklärung disulfidverbrückter Peptide und Proteine erschwert.

Das Ziel dieser Arbeit war es, die Faltung von kleinen, mehrfach disulfidverbrückten Peptide unter verschiedenen Bedingungen zu analysieren und Struktur-Aktivitätsuntersuchungen der erhaltenen Produkte durchzuführen. Im Rahmen dieses Projektes wurden unterschiedliche Conotoxine mittels Festphasenpeptidsynthese hergestellt. Zur Ausbildung der Disulfidbrücken wurden verschiedene Methoden eingesetzt: oxidative Selbstfaltung, Oxidation in ionischen Flüssigkeiten und stufenweise Oxidation in Verbindung mit einer orthogonalen Schutzgruppenstrategie. Die oxidierten Conotoxine wurden anschließend funktionell untersucht und durch verschiedene analytische Methoden charakterisiert. Mittels NMR-Spektroskopie wurden die dreidimensionalen Strukturen der erhaltenen Peptide aufgeklärt. Weiterhin konnte nach partieller Reduktion und Derivatisierung die Disulfidverbrückung durch verschiedene massenspektrometrische Verfahren (LC-ESI und MALDI) bestimmt werden. Die Ergebnisse zeigen die jeweiligen Vorteile und Limitierungen der verwendeten analytischen Methoden hinsichtlich der Unterscheidung verschiedener disulfidverbrückter Conotoxinisomere in Abhängigkeit von ihrer primären Aminosäuresequenz. Parallel zu den biologischen Testungen der hergestellten Peptide wurden computerbasierte Untersuchungen durchgeführt um ihre Struktur-Aktivitätsbeziehungen zu analysieren. Die vorgestellte Dissertation erweitert das Wissen über Conotoxine und die Bildung von Disulfidbrücken mit ihren Auswirkungen auf die biologische Aktivität und demonstriert darüber hinaus die schwierige Strukturaufklärung und Charakterisierung.

Content

1. Introduction	1
2. Theoretical background	2
2.1 Drug lead substances.....	2
2.1.1 Venom for drugs	2
2.1.2 Peptides as drugs	3
2.2 Cysteine-rich peptides.....	5
2.2.1 Conotoxins	5
2.2.2 Voltage-gated ion channels.....	11
2.3. Synthesis of cysteine-rich peptides.....	14
2.3.1. Protecting group strategies	17
2.3.2. Oxidative folding	18
2.3.3. Room temperature Ionic Liquids for peptide oxidation	19
2.4 Analytical methods for multiple disulfide-bridged peptides.....	20
2.4.1 LC-ESI mass spectrometry	23
2.4.2 MALDI mass spectrometry	24
2.4.3 Two-dimensional NMR spectroscopy	26
3. Thesis outline	28
4. Manuscripts	31
4.1 Ionic Liquid applications in peptide chemistry: synthesis, purification and analytical characterization processes	31
4.2. Ionic Liquids as reaction media for oxidative folding and native chemical ligation of cysteine-containing peptides.....	61
4.3 Application of room-temperature aprotic and protic Ionic Liquids for oxidative folding of cysteine-rich peptides	67
4.4 Molecular interaction of δ -conopeptide EVIA with voltage-gated Na ⁺ channels	81

Content

4.5 Computational study on μ -conotoxin conformations using molecular dynamics and computer visualization methods	94
4.6 Conformational peptide isomers revisited: The impact of disulfide connectivity on structure and bioactivity	105
5. Conclusions.....	118
5.1 Oxidative folding of cysteine-rich peptides in Ionic Liquids	118
5.2 Analytical characterization and structure-activity relationships of conotoxins	119
6. Abbreviations	124
7. List of figures.....	126
8. List of tables	127
9. References.....	128
Appendix.....	140

1. Introduction

Over the last decades peptides were established in the drug market and filled the gap between small-molecules and larger biopharmaceuticals.^{1–3} A rich source of therapeutically useful peptides is the venom of different animals, for example snakes, scorpions, spiders, and cone snails, which has been evolved over millions of years.⁴ One group of toxins of particular interest are conotoxins, which are small, disulfide-rich peptides acting on different biological targets such as ion channels or receptors, isolated from cone snails of the genus *Conus*.⁵ The venom contains a complex cocktail of bioactive substances which is used for self-defence or hunting prey.⁶ With over 700 species known, these marine invertebrates possess large potential for identifying novel pharmacologically active molecules, especially for the treatment of pain.⁵ One conotoxin, Ziconotide (Prialt®), that is acting on calcium-channels, is already available to treat severe chronic pain.⁷ Conotoxins are classified according to their pharmacological target, for instance μ -conotoxins block voltage-gated sodium channels with high selectivity and affinity.⁵ Therefore, several conotoxins represent a promising tool for drug development.^{1,8} Since the snail produces only small amounts of the venom, there is a big need for chemical synthesis of larger amounts after identification of the interesting components.^{5,9} Although standard solid phase peptide synthesis is established, preparation and characterization of cysteine-containing peptides forming disulfide bonds is still challenging.^{1,5} In particular the selective formation of disulfide isomers and possible rearrangement of disulfide bridges during peptide preparation is difficult essentially complicating synthesis and elucidation of cysteine-rich peptides and proteins.^{1,5} In addition, disulfide bonds are the most common posttranslational modifications in peptides and proteins stabilizing and conserving the bioactive conformation of the molecules.¹⁰

In this work, formation of disulfide bonds during the preparation of selected conotoxins is analyzed using a set of methods for structure elucidation. The synthesized peptides are tested for their biological activity on voltage-gated sodium channels. In addition, *in silico* studies contribute to structure-activity relationship studies.

2. Theoretical background

2.1 Drug lead substances

2.1.1 Venom for drugs

The father of toxicology Paracelsus (Theophrastus Bombastus von Hohenheim) established the phrase: "The dose makes the poison" centuries ago.¹¹ This statement still holds true today, especially if considering the most efficient and selective compounds in the process of drug discovery.¹² Natural products remain an inexhaustible source for the discovery of novel therapeutics.¹³ The venom of different animals, which has evolved over millions of years, represents a rich source of natural compounds and contains a complex mixture of bioactive substances.¹⁴ Besides non-peptidic toxins such as tetrodotoxin (TTX), saxitoxin or batrachotoxin typically developed for self-defence, peptide toxins are usually found in species with potential of envenomation. Animals like snakes, scorpions, fish or cone snails produce a species-specific venom mixture containing compounds with convergent recruited scaffolds like cystatins, cysteine-rich secretory proteins (CRISPs) or defensins.^{15,16} The advantages of using venom-derived peptides are discussed in detail in chapter 2.1.2. Today, numerous of these peptides are commercially available on the market or are in various stages of clinical or preclinical development.¹⁷ Several companies have optimized technologies for venom screening assays, venom libraries, hit-to-lead identification (e.g. *de novo* sequencing) and lead optimization.^{18,19} These "venomics" approaches have meanwhile been established as an important part of modern drug discovery with large potential of identifying novel pharmacologically active molecules.⁴ For the characterization of low-molecular weight organic compounds of a venom cocktail, proper application and sustained progress of analytical methods is pivotal (chapter 2.4). The most frequently addressed biological targets of toxins comprise different isoforms of ion channels, G-protein-coupled receptors (GPCRs) and other receptors such as nicotinic acetylcholine receptors (nAChRs).⁴ The development of the angiotensin-converting enzyme (ACE) inhibitor Captopril for the treatment of hypertension is mentioned as an example for drug development from venom. Initially, the nonapeptide Teprotide derived from the Brazilian pit viper *Bothrops jararaca* was chosen as lead structure.^{16,20} The minimal substrate-binding site was identified to be the tripeptide Phe-Ala-Pro using the crystal structure of carboxypeptidase A139, although later the ACE-Lisinopril complex crystal structure revealed low similarity between both enzymes.

Several other examples for the successful application of venomics support this method as a new paradigm for drug discovery as is described in more detail in the following chapter.^{4,14}

2.1.2 Peptides as drugs

Meanwhile, significant improvements in peptide synthesis protocols and analytical characterization processes (chapters 2.3 and 2.4) led to the introduction of peptides in the pharmaceutical market. Various biomedical applications, such as tissue engineering for regenerative medicine, vaccine production, medical imaging technologies, and drug delivery, nowadays employ peptides.¹ However, peptides are often of limited use due to a rather low oral bioavailability, which is usually circumvented by alternative administration strategies, e.g. by injection (subcutaneous, intravenous or intrathecal).²¹ In addition, peptides lack stability against enzymatic degradation. Approaches such as backbone cyclization, introduction of *D*-amino acids or replacing cysteine by selenocysteine can increase the peptide stability and therefore lead to an improvement of the delivery methods (e.g. inhalation, transdermal, intranasal).²¹ Formation of multiple disulfide bonds represents another modification and result in a distinct topology and conformation. Such constrained peptides combine the unique properties of small molecules and biologics and consequently fill the gap between both (Figure 1).¹

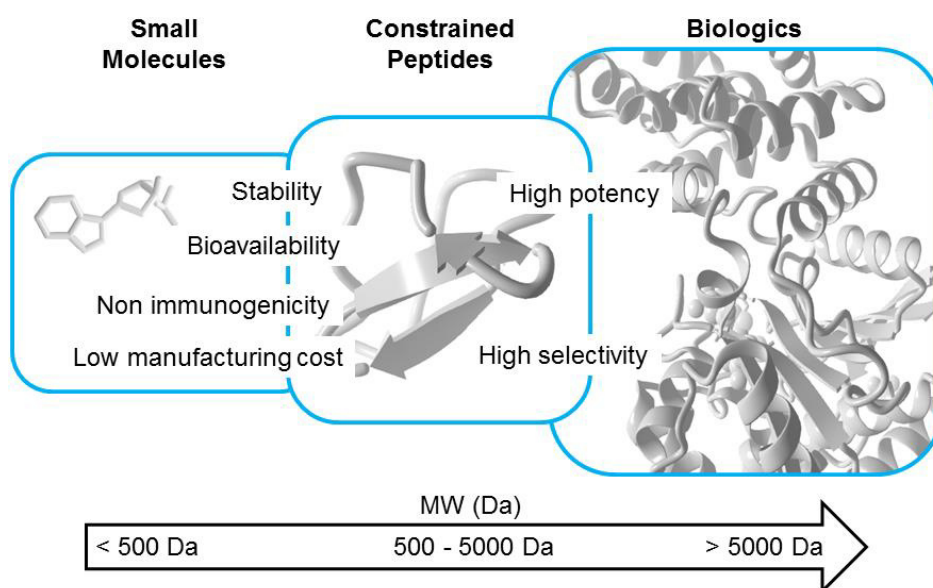


Figure 1. Advantage of constrained peptides as therapeutics, filling the gap between small molecules and biologics by combining the properties of each type of drugs. Modified from 1

Disulfide-bridged peptides do not only appear in the venomous animals mentioned above, but also in plants, as can be exemplified with the cyclotide kalata B1 which endures high temperatures and enzymatic degradation in its oxidized form.²² This peptide, derived from the plant *Oldenlandia affinis*, remains its active conformation and tolerates harsh conditions during tea preparation and oral application because of multiple disulfide bonds. Therefore, a

2. Theoretical background

further application of such a disulfide-bond framework represents the molecular grafting combining sequence and conformation characteristics from different bioactive peptides.^{22,23}

Several advantages of constrained peptides are higher stability against enzymatic degradation, improved bioavailability, lower costs for large-scale production compared to biologics, together with higher potency, affinity, selectivity and fewer side effects compared to small molecules.¹ Therefore, peptides were introduced in the market with high global sales in 2007 as can be exemplified with Copaxone® (\$ 3.33 billions), Lupron® (\$ 1.88 billions), Byetta® (\$ 967 millions), and Forteo® (\$ 709 millions). Three years later, over 50 peptide therapeutics have been approved for the market.^{2,24} A large number of venom-derived peptides occur within these peptides. For instance, Exenatide (Byetta®), a 39 amino acids peptide, derived from the saliva of the venomous Gila monster *Heloderma suspectum*, was approved for the treatment of type 2 diabetes mellitus as the first glucagon-like peptide-1 (GLP-1) receptor agonist.²⁵

Peptide production is performed by a variety of methods including chemical and biological synthesis. Beginning with solution-phase synthesis (SPS) introduced in 1953 by du Vigneaud peptide drugs became accessible for commercial usage. However, the method is less efficient considering time, labour and cost economics.^{2,26} A milestone in peptide synthesis overcoming the previous limitations was achieved by Merrifield in 1963 by the introduction of solid-phase peptide synthesis (SPPS).²⁷ In this approach, an insoluble polymer resin is utilized for immobilization of primarily *tert*-butyloxycarbonyl (Boc)- or 9-fluorenylmethyloxycarbonyl (Fmoc)-protected amino acids, depending on the α -amino acid protecting group (PG) strategy applied. Alternating steps of acid- (in case of Boc) or base-labile (in case of Fmoc) deprotection and coupling of the following amino acid with coupling reagents, e.g. 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) is performed together with extensive washing steps. Various protecting groups have been introduced for peptide preparation using different strategies.²⁸ In recent years, however, the development of alternative approaches which facilitate complex peptide synthesis, such as the combination of SPS and SPPS (hybride synthesis)²⁹ or chemoselective ligation strategies like native chemical ligation (NCL),³⁰ have been promoted. The latter was established already in 1953 by Wieland *et al.* by combining two peptides via a thioester and a cysteine residue in a rearrangement reaction.³¹ After several decades and respective progress in SPPS, Dawson and Kent extended this principle in 1994 for cysteine-containing peptides.³² With this simple technique they were able to synthesize human interleukin 8 (72 AA) by ligation of 33mer and 39mer peptide fragments. Since then even longer peptides and small proteins were prepared using one or more ligation steps.³³ Furthermore, the limitations of cysteine containing sequences have been overcome with different ligation methods such

as methionine, histidine or glycine ligation.³⁴ Among the proteinogenic amino acids cysteine is unique since it is able to introduce rigid structures in peptides through disulfide bond formation by connecting two cysteine residues in an intra- or intermolecular way.^{1,35} This special feature of cysteine will be discussed in more detail in chapter 2.3.1.

Besides the chemical approach processes like enzymatic synthesis, native isolation as well as fermentation and recombinant expression have been established for bioactive peptide production. Although these methods are sometimes of limited use due to elaborate purification procedures or the incorporation of non-natural amino acids, several systems were established employing these approaches, e.g. the previously mentioned Forteo® was expressed in *Escherichia coli*.^{1,2,29}

2.2 Cysteine-rich peptides

2.2.1 Conotoxins

60 years ago A.J. Kohn described the marine cone snail of the genus *Conus* for the first time. The cone snail actively hunts fish using a harpoon and secreting a powerful neurotoxin.³⁶ As a consequence of this observation the ecological niche of cone snails was changed from carrion eaters to active predators producing toxic venom.³⁶ The lethal potential of the toxin mixture was discovered when the venom was accidentally injected into humans by a cone snail for self-defence. These people were collecting the beautiful shell which fascinated human cultures from ancient times up to now (Figure 2).^{37,38} The aforementioned toxin consists of neuroactive peptides, also called conopeptides, which are derived from marine cone snails.³⁹ Cone snails comprise marine invertebrates of hunting gastropods with more than 700 species known so far.^{39,40} Because of the large diversity Puillandre *et al.* proposed a new classification in 2015 according to molecular phylogenetic analysis dividing the family (*Conidae*) into four genera, which is in contrast to the former system established by Cruz *et al.* in 1985.^{41,42}



Figure 2. Shells from different fish-hunting cone snails: *Conus ermineus* (left), *Conus purpurascens* (middle) and *Conus striatus* (right). Modified from 68,69

2. Theoretical background

Nevertheless, the venom of each species is reported to contain 50 to 200 different components representing a highly complex cocktail of various peptides and proteins potentially binding to different biological targets.³⁸ Today, due to the efforts made in the field of analytics, especially mass spectrometry, up to >2500 masses can be detected in a single cone snail species venom depending on factors which are described below.⁴³ Thus, more than >70.000 pharmacologically relevant compounds are potential drug leads derived from cone snails, whereof only ~ 2000 toxins have been identified so far.⁴⁴

The diversity of cone snail venom and its composition depends on several factors. First of all, different cone snail species produce miscellaneous variations of venom for hunting prey, self-defence or to deter competitors.⁴⁵ Besides, the toxin is adapted to the respective dietary, i.e. worms (vermivorous), molluscs (molluscivorous) or fish (piscivorous).³⁹ Furthermore, the components differ depending on the geographical regions and also individual variations can be found within a species.^{46,47} These aspects indicate that each individual specimen may adjust and brew its own specific toxin cocktail.⁴⁸

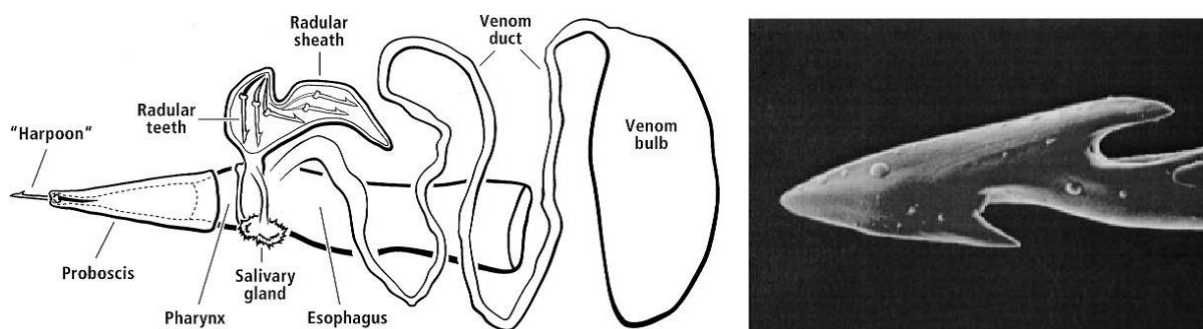


Figure 3. (Left) Schematic view of a cone snail venomous apparatus. (Right) Harpoon-like radular teeth enlarged by scanning electron micrograph from *Conus purpurascens*. Modified from 45

The toxin envenomation system consists of a complex venom apparatus shown in Figure 3.⁴⁵ The complete mechanism of venom production is still unclear and further research is needed to enlighten the complex procedure. One difficulty arises from the observation that the venom composition highly differs from different levels (e.g. injected venom, venom duct) obtained. In general the toxin cocktail is produced in the venom duct and transferred for injection through contraction of the muscular venom bulb. The toxin is finally injected with a radular tooth acting as a hollow harpoon that is produced and stored in the radular sheath. From there the harpoon is transferred through the pharynx to the proboscis, where a single tooth is prepared for injection of the venom. Upon toxin application, the snail needs regeneration time to recover full toxin potential.^{45,49,50}

First analysis of the venom mixture was made in 1977 by Spence *et al.* separating the crude poison of *Conus geographus* into several fractions. The following application into different

tissues revealed the specific action.⁵¹ Further detailed isolation and characterization of the toxin mixture of *C. geographus* and *C. magus* was performed in the laboratories of B. M. Olivera and L. J. Cruz later on.^{52,53} These piscivorous species were chosen because their venom contains more active components against vertebrates compared to species hunting invertebrates. Starting from this point, the analytical methods evolved rapidly. Improved purification and identification of the bioactive compounds of the toxin cocktail facilitate characterization nowadays. Therefore, it is now possible to describe the individual fractions and their respective biological activity more precisely (Figure 4). Distinct toxin fractions were thus identified as small, disulfide-rich peptides (10-30 AS). The peptides are characterized by high structural conservation, numerous post-translational modifications and, in general, were found to be smaller than most other toxins derived from animals like snakes, scorpions or spiders.^{6,49,53}

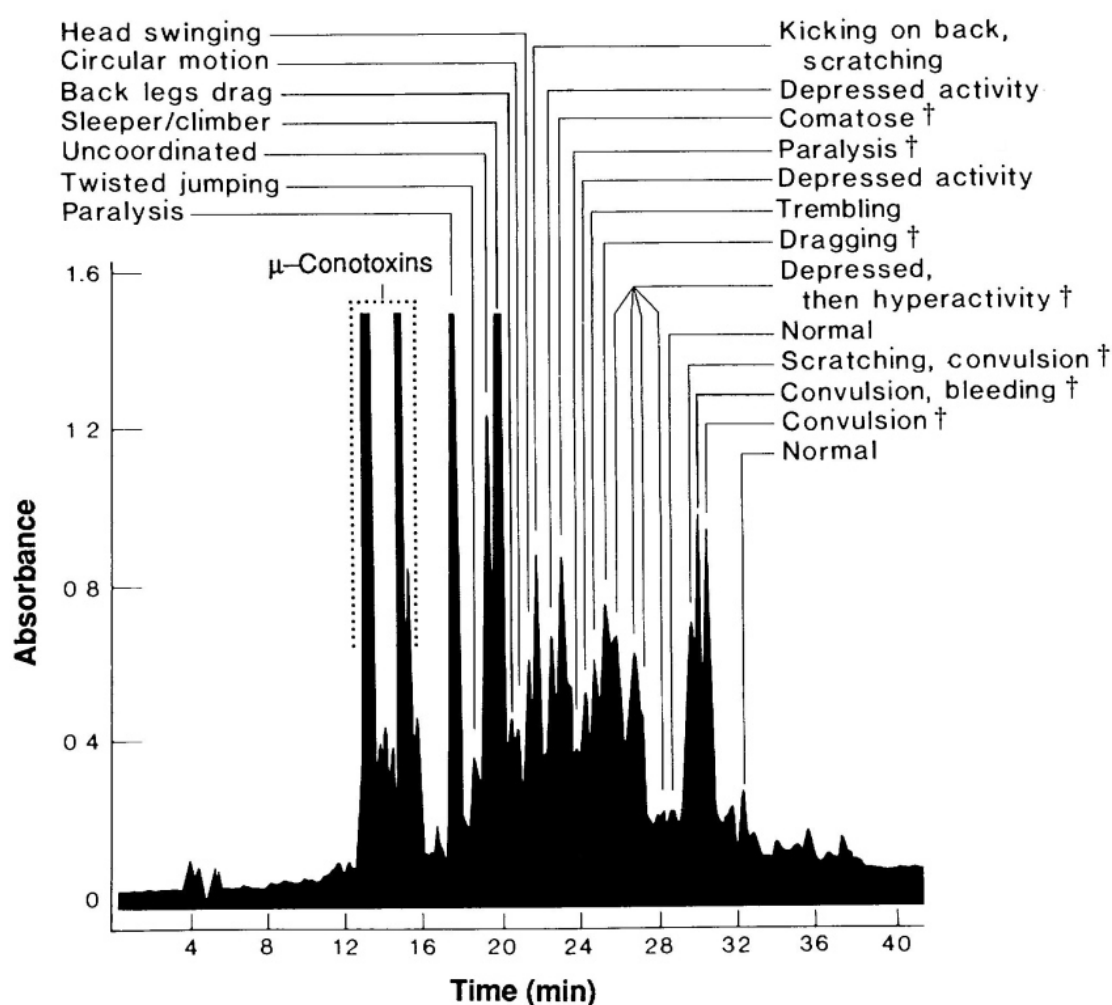


Figure 4. Crude venom of *Conus geographus* purified with HPLC and fractions labelled with respective biological activity. This was observed by intracranially injection of ~0.5 – 2 nmol of each fraction into mice. The fractions with † caused death at in least one injected animal. Modified from 49

2. Theoretical background

Originally, the so-called conopeptides derived from cone snails were divided into two groups. The first group is composed of the disulfide rich peptides (conotoxins), defined for peptides with more than two cysteine residues. In contrast, the second group consists of the non-disulfide rich peptides with several subclasses such as e.g. contulakin, conopressin or contryphan with two or less cysteines.³⁸ However, this discrimination became redundant because phylogenetic analysis indicated similar signal sequences within the conopeptides previously classified according to the cysteine content.⁵⁴ Classification of conotoxins has come a long way starting from a strict ordered concept that is nowadays inappropriate due to several novel insights. For example subtypes of classification were established based on the genetic superfamily, the cysteine framework, the biological target and the conotoxin family.³⁸ In the respective publication by Terlau and Olivera from 2004 only seven superfamilies have been reported.³⁸ However, nowadays more than 30 superfamilies were described and several new superfamilies are permanently reported in current publications.⁴⁴ Some of these families have not been fully described yet, leaving some subtypes of classification like biological target open, which are necessary for complete characterization.⁴⁴ Today, new conotoxins are usually identified using cDNA sequencing and resulting in gene superfamilies as classification method established by Woodward *et al.* in 1990.⁵⁵ The complex relationship between the genetic superfamily, the pharmacological family, the cysteine framework, and the number of cysteine residues is described in Figure 5 with the respective connections within the subgroups.^{38,44,56–58} Conotoxin classifications important for this work are highlighted in bold.

The first isolated and described conotoxin was α -GI containing two disulfide bridges and belonging to the A-superfamily.^{6,59} Greek symbol α indicates the respective pharmacological family and the biological target, here nicotinic acetylcholine receptors (nAChRs), and the “G” refers to peptides isolated from the genus *Conus geographus*. The roman number “I” represents the cysteine framework with the corresponding connectivity, here CC-C-C containing bridges Cys1-Cys3 and Cys2-Cys4.^{59,60} The last Arabic letter in the conotoxin name (e.g. μ -GIIIA, μ -GIIIB, μ -GIIC) represents the first, second or third peptide discovered with the same disulfide framework and the species origin.^{45,61} Although conotoxins with a distinct cysteine-framework share an identical cysteine scaffold, the residues within the loops can highly differ in their sequence. This indicates a large variation for each residue with the exception for those conserved amino acids essential for bioactivity.⁶² For instance, all μ -conotoxins possess the same cysteine-framework (Cys1-Cys4, Cys2-Cys5, Cys3-Cys6), but the resulting structures, target affinity and selectivity vary highly from each other.⁶³ Therefore, elucidation the relationship of the primary amino acid sequence and the biological effect of μ -conotoxins is an interesting research target. In case of the μ -conotoxins the respective

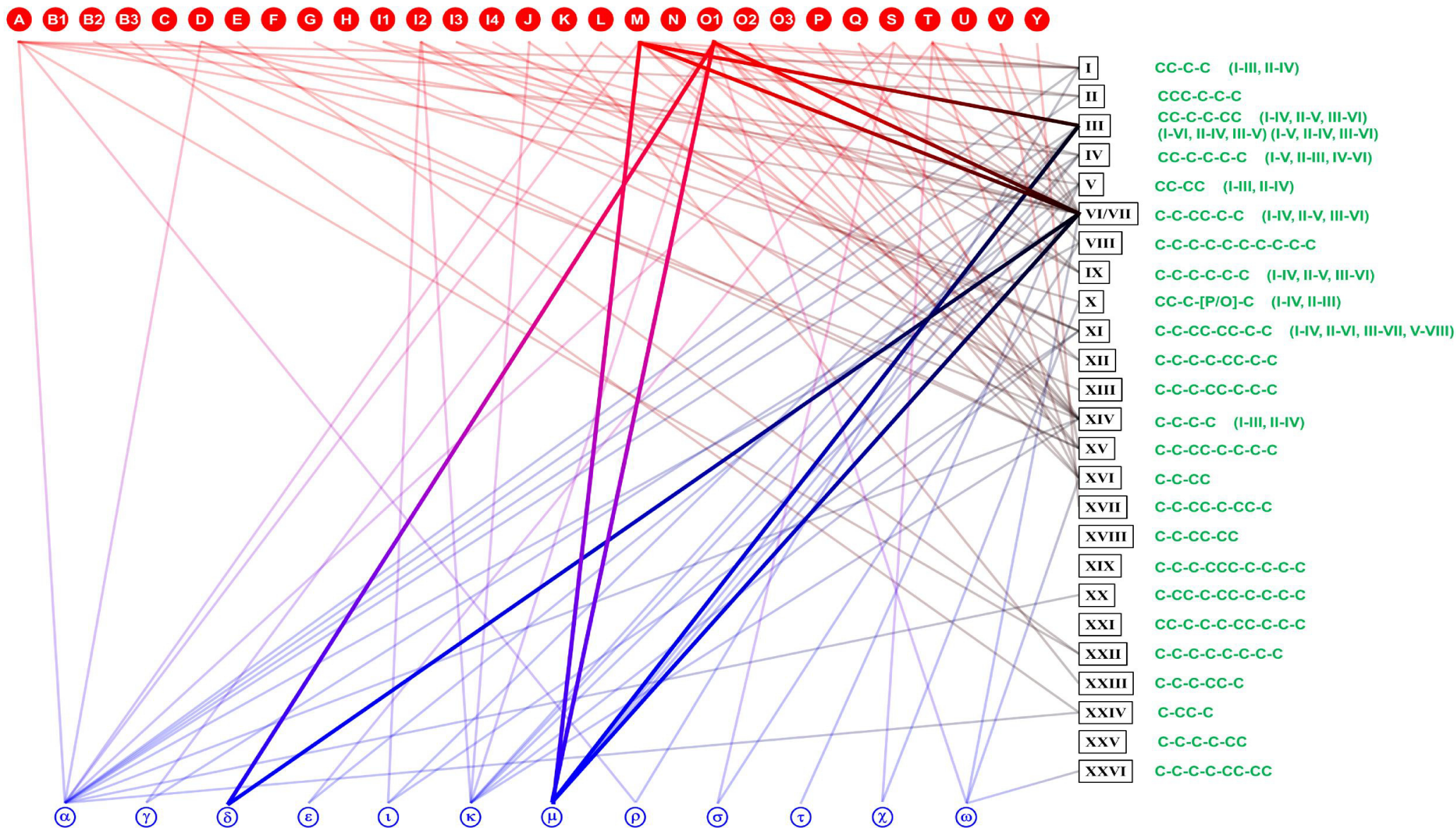


Figure 5. Conotoxin classification displaying the relationship between genetic superfamily (red), pharmacological family (blue), cysteine framework (black) and cysteines included with respective connectivity if known (green). Modified from 56–58

2. Theoretical background

biological target are the TTX-sensitive/resistant voltage-gated sodium channels (VGSCs) (see chapter 2.2.2.).^{63,64}

The M-superfamily, for example, is divided into five subfamilies based on the amino acid residues in the third loop region. Members possess primarily six cysteine residues with the disulfide framework III and two adjacent cysteine pairs (CC-C-C-CC).^{65,66} These superfamilies contain μ -conotoxins (target VGSCs), κ M-conotoxins (target voltage-gated K⁺ channels) and ψ -conotoxins (target nAChRs). Most of the μ -conotoxins belong to the M4/5-subfamily, e.g. μ -PIIIA from *C. purpurascens*, μ -SIIIA from *C. striatus* or μ -GIIIA from *C. geographus*, always the first members discovered.^{56,66} Another example is δ -EVIA, which was identified from *C. ermineus*, bearing cysteine framework VI (C-C-CC-C-C) and that blocks the inactivation of VGSCs. EVIA is part of the O-superfamily, which is divided into three subfamilies.^{56,67} Formation and also elucidation of the respective disulfide connectivity will be discussed in more detail in the following (chapter 2.3.1). Since 2008 Q. Kaas developed and perpetually updated an internet database, *conoserver*, with all relevant information about each conotoxin, their respective origin, biological activity and references.^{68,69}

The venom of each cone snail species can have different physiological effects depending on the tissue and therefore results in a huge therapeutic potential as previously mentioned.^{70,71} Since 2004, the conotoxin Ziconotide (Prialt®) is commercially available on the market for pain treatment. Ziconotide or ω -MVIIA is a 25mer peptide which was derived from *Conus magus* as was first described in 1985 by Olivera *et al.*^{6,72} It selectively inhibits voltage-dependent N-type calcium channel (Ca_v2.2), therefore exhibits a therapeutic potential for the treatment of chronic pain, and also for opiate-resistant patients.^{7,55} Although Ziconotide is at least 10-times more potent compared to a common analgesic like morphine, its application is limited due to several side effects like for example ataxia, abnormal gait, dizziness, nystagmus and somnolence.^{55,77} For this reason widespread use is problematic. The application form, i.e. intrathecal infusion via a permanent implanted pump is restricted due to the necessity of a surgery. Nevertheless, Ziconotide was shown to help people suffering especially from cancer- or AIDS-related chronic pain enabling higher quality of life.^{7,52,73}

Further conotoxins are tested in clinical trials, with the challenge transferring the beneficial use from animal models to human application and finally a profitable drug. However, several reasons led to a failure of the respective studies. The ω -conotoxin CVID (Leconotide, AM336 or CNSB0047) from *Conus catus* effects Ca_v2.2 channel like Ziconotide with the advantage of a higher selectivity and less side effects by intrathecal and intravenous application to rats.⁷⁴ Clinical trials were stopped due to the liquidation of the developing company.¹⁷ A further example is the α -conotoxin Vc1.1 (ACV 1) from *Conus victoriae* that decreases

calcium ion currents in nociceptive neurons. Surprisingly, the effect appears to be mediated by an altered mechanism, the conotoxin inhibits the $\text{Ca}_v2.2$ channel indirectly and voltage-independently by acting on G protein-coupled γ -Aminobutyric acid (GABA_B) receptor causing unwanted side effects. Furthermore, oral availability of this compound was reached after head-to-tail cyclization including a six residues linker sequence increasing the stability towards degradation *in vivo*. Nevertheless, further biological testing is required to clarify the mechanism and detailed mode of action and to avoid unexpected side effects, lack of efficacy or dose dependent effects.^{17,75–77}

Focussing on the therapeutic potential of μ - and δ -conotoxins there are several advantages and limitations for possible drug targets acting on VGSC.⁷⁸ In general μ -conotoxins affect tetrodotoxin (TTX)-sensitive VGSC with the highest affinity against skeletal muscle isoform $\text{Na}_v1.4$ and/or $\text{Na}_v1.2$.^{79,80} A more detailed description about specificity of selected peptides is included in chapter 2.2.2. However, medication-dependent properties reveal potent analgesic activity by intrathecal application in animal model, whereas intravenous injection generally causes paralysis and death, restricting the drug potential.⁷⁹ On the other hand δ -conotoxins were considered to have a lower therapeutic potential because they block the inactivation of VGSCs. Furthermore, the required elucidation of structure-activity relationships are rare because of challenging handling of the toxins during synthesis, folding and purification of these highly hydrophobic peptides.⁷⁹

Another import aspect of conotoxins is their target specificity within subtypes of ion channels. The toxin selectively blocks distinct ion channel isoforms and simultaneously remains others untouched (chapter 2.2.2). Therefore, they are applied tools for discrimination between different ion channels and there subtypes. Although conotoxins are available on the market from several companies (e.g. Bachem[®], Smartox[®], Iris Biotech GmbH[®]), their synthesis as well as their structural determination and elucidation of the formed disulfide bridging is still challenging.^{1,2}

2.2.2 Voltage-gated ion channels

Contiguous to nAChRs, voltage gated ion channels are one main target of conotoxins and other animal toxins. These channels are membrane proteins that transfer inorganic ions like Na^+ , K^+ or Ca^{2+} depending on their state (open or closed). These channels exist in most vertebrates and invertebrates.⁸¹ Ion channels are essential in propagation of action potential in excitable cells like neurons or muscle cells.⁸² Starting in 1952 Hodgkin and Huxley established the so-called voltage clamp method that allows to quantify the membrane current in such cells.⁸³ Although several species possess various ion channels in different tissues,

2. Theoretical background

evolution developed nine human isoforms of voltage gated sodium channels, which were characterized so far (Table 1). The channels can be expressed in different tissues such as peripheral and central nervous system, skeletal muscle or heart and have different resistances against TTX, a toxin isolated from pufferfish, that is utilized as a pharmacological standard for differentiating the channel subtypes.⁸¹ In addition other selectively acting components of the toxins (e.g. snake-, scorpion- or conotoxins) can be utilized to characterise the structure and function of ion channels.⁸⁴ VGSC consists of a α - and one or more β -subunits depending on the specific channel subtype. An active channel is usually established by the pivotal α -subunit. This consists of four homologous domains with each domain build by six transmembrane segments (S1-6) (Figure 6). Segments S1-4 built the voltage sensor domain and S5-6 form the channel pore.⁸⁵ Although the complete mechanism of action of the channel is still unknown, some details have been identified. The voltage sensor domain S4 induces a conformational change upon depolarization that leads to opening of the channel pore. This process is usually followed by intracellular fast inactivation, milliseconds after activation. Here, the loop connection between domain 3 and 4 (D3,4) blocks the pore from intracellular space with the residues IFM (Figure 6).⁸⁶

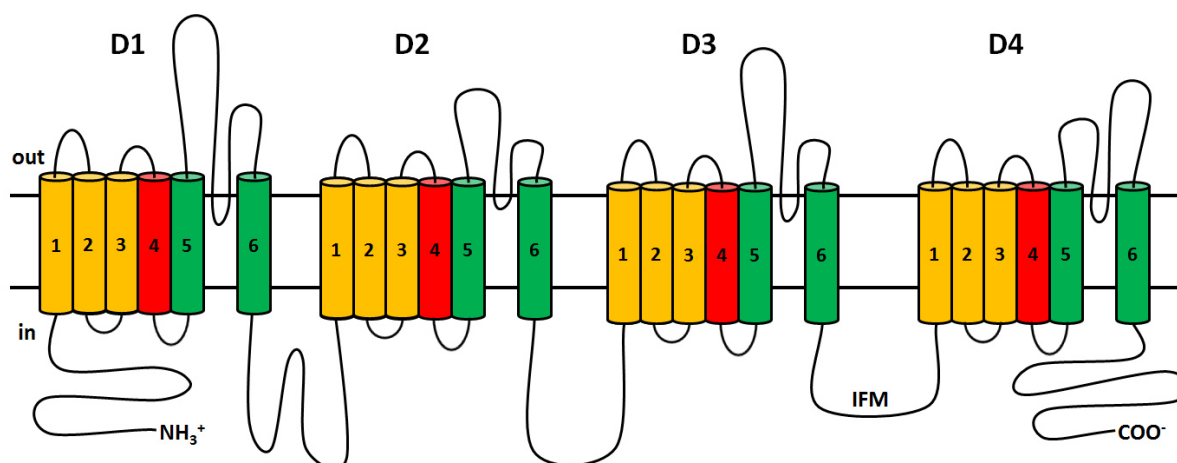


Figure 6. Voltage gated sodium channel α -subunit consisting of 4 homologous domains (D1-4) with each consisting of 6 transmembrane segments (S1-6). The voltage sensor S4 is marked in red and the channel forming units S5/6 are shown in green. The loop responsible for fast inactivation between D3 and D4 is labelled with the relevant residues IFM relevant for fast inactivation. Modified from 86

Targeting this transfer of action potential for distinct pain-mediating subtypes of VGSCs specifically, results in powerful analgesics. Compared to small molecules like morphine or lidocaine, conotoxins have a higher selectivity for the different VGSC subtypes.⁸⁷ This specificity is summarized for some μ -conotoxins and the respective channel in Table 1. Indeed, distinct subtypes play a pivotal role in pain treatment such as e.g. Na_v1.7, 1.8 or 1.9. These represent interesting drug leads for toxins that block these channels selectively.⁸⁸

channel	TTX resistance ⁸⁷	Distribution (human) ^{81,87}	μ-KIIIA	μ-GIIIA	μ-PIIIA	μ-SIIIA	μ-SmIIIA
Na_v1.1	Sensitive	CNS, PNS	290 nM +/-110 ^{89#}	260 nM +/-90 ^{90#}	120 nM +/-13 ⁹⁰	11 μM +/-1.2 ⁹¹	3.8 nM +/-1.3 ^{90#}
Na_v1.2	Sensitive	CNS	61 nM +/-5, ⁹² 5 nM ⁹³	17.8 μM +/-1.7, ⁹⁰ 2.7 μM ⁹⁴	620 nM +/-70, ⁹⁰ 690 nM ⁹⁴	790 nM ⁹⁵	1.3 nM +/-0.7 ^{90#}
Na_v1.3	Sensitive	CNS, Foetal DRG	3.6 μM +/-0.3, ⁹² 8 μM +/-1 ⁸⁹	>10 μM ⁹⁰	3.2 μM +/-0.8 ⁹⁰	11 μM +/-1.4, ⁹¹ 5.3 μM ⁹⁵	35 nM +/-14 ^{90#}
Na_v1.4	Sensitive	Skeletal muscle	48 nM +/-6, ⁹² 54 nM, ⁹³ 90 nM +/-17 ⁸⁹	19 nM+/-2, ⁹⁰ 19nM ⁹⁴	36 nM +/-8, ⁹⁰ 41 nM, ⁹⁴ 47-204 nM +/-7-25* ⁹⁶	130 nM +/-15, ⁹¹ 330 nM, ⁹⁵ 560 nM +/-290 ⁹⁷	0.22 nM +/-0.1 ^{90#}
Na_v1.5	Resistant	Heart muscle	287 μM +/-40, ⁸⁹ >10 μM ⁹² <i>H. sapiens</i>	>10 μM ⁹⁰	>10 μM ⁹⁰	251 μM +/-38 ⁹¹	1.3 μM +/-0.25 ⁹⁰
Na_v1.6**	Sensitive	PNS, CNS	183 nM +/-31, ⁹² 240 nM +/- 30 ⁹⁰	680 nM+/-60 ⁹⁰	100 nM+/-20 ⁹⁰	760 nM +/-60 ⁹¹	160 nM +/-40 ⁹⁰ <i>R. norvegicus</i>
Na_v1.7	Sensitive	PNS, schwann cells	147 nM ⁹³	>10 μM, ⁹⁰ 6 μM ⁹⁴	>10 μM, ⁹⁰ 6.2 μM, ⁹⁴ <i>H. sapiens</i> 3.1 μM ⁹⁴	65 μM +/-6 ⁹¹	1.3 μM +/-0.23 ⁹⁰
Na_v1.8	Resistant	DRG	>10 μM ⁹² <i>H. sapiens</i>	>10 μM ⁹⁰	>10 μM ⁹⁰	>10 μM ⁹⁰	>10 μM ⁹⁰
Na_v1.9	Resistant	PNS	-	-	-	-	-

Table 1. Subtype specificity of different μ-conotoxins on VGSCs relating to their IC50 values. # Only K_D available. CNS = central nervous system, DRG = dorsal root ganglia, PNS = peripheral nervous system; VGSC from *R. norvegicus* was used in general unless otherwise mentioned; * indicates bioactivity for respective isomer isolated; **1.6 from *M. musculus*.

2. Theoretical background

Mutations of VGSC can cause several symptoms or diseases, e.g. selective deficiencies in pain pathways initiated by Na_v1.8 channel knockouts or in the Na_v1.9 channel.^{87,98} Exact channel binding sites of a toxin are identified by applying a highly specific blocking toxin to different subtype chimeras of two VGSCs. From these, one channel is blocked while the other is unaffected by the toxin. The chimeras are mutated versions with single domains or segments of the respective channels replaced, indicating the relevance for binding and blocking of the toxin. In addition, single point mutants reveal more detailed information about the interacting positions.⁹⁷

Nowadays, drug design is also performed *in silico* by predicting the interaction between channel and channel modifier and therefore giving a hint on the respective bioactivity of the applied substance.⁹⁹ Essential for these studies are structural data from both sides: the VGSC, where only little data is available from x-ray structures, and the modulator. Structural data of the modulator can be derived from spectroscopic methods such as NMR (chapter 2.4.3). If there is no structural data existing from a distinct subtype, a homology model can be created from a similar subtype with available structural data.⁹⁹ To simplify these models parts of the channel (unimportant for the simulated interaction) are often excluded or exchanged with the residues present in the desired isoform.¹⁰⁰ These results, together with the electrophysiology measurements, can be utilized to understand the interaction of the channel and the channel modulator.⁹⁹ In the last two years several computational studies prove the rising interest in binding and blocking of μ -conotoxins to VGSC.^{100–102}

2.3. Synthesis of cysteine-rich peptides

Within the field of peptide synthesis preparation of cysteine-rich peptides is a challenging task especially the oxidation and selective formation of disulfide bridges.¹ Nevertheless, enormous research interest and therapeutic potential of cysteine-rich peptides make large scale synthesis of these peptides essential. In the past decades the knowledge of synthesis and oxidation of cysteine containing peptides has expanded due to the work of G. Barany, L. Moroder and F. Albericio.^{1,35,103–105} In general, three different approaches were established for synthesis of disulfide containing peptides.¹⁰⁶ First of all, folding of the disulfide bridging on solid support is possible but requires low resin loading for preferred intramolecular oxidation and appropriate orthogonal protecting groups. During the second strategy the first disulfide bond is formed on solid phase. This restricts the conformational freedom for the formation of further disulfide bridges in solution after cleavage of the partially oxidized peptide. Finally, the formation of disulfide bridges can be performed completely in solution with different approaches e.g. oxidative folding forming all bridges with one step or successive oxidation applying suitable PG strategy (chapters 2.3.1 to 2.3.3).¹⁰⁶ During oxidation of the cysteines to

respective disulfide connectivity, formation of different isomers by rearrangement is a big issue.³⁵

In 1973 Anfinsen established the dogma that polypeptide chains or proteins fold into a thermodynamically preferred “native” isomer based on the example of ribonuclease A. Furthermore, he stated that all other isomeric forms which appeared during preparation had lower biological activity.^{107,108} Extensive research elucidating the folding mechanism of disulfide containing peptides found in several molecule classes such as hormones, growth factors, enzyme inhibitors, neurotransmitters, toxins, and antimicrobial peptides to particular cysteine knot motives was necessary to support Anfinsen’s postulate.^{35,109–111} The characterization of the specific folding pathway of individual peptides like e.g. insulin, bovine pancreatic trypsin inhibitor (BPTI) and hirudin revealed different transition states for the respective peptide.¹⁰⁹ However, non-native isomer formation during various folding procedures is highly sequence and parameter dependent and behaviour of undescribed and uncharacterized bioactive peptides is difficult to predict.^{108,109,112} This problem is described in the context of conotoxins in more detail in the following chapter with respective examples.

Today for the α -conotoxins with only two disulfide bridges resulting in three possible isomers synthesis, structure elucidation and biological activity is established and in focus of recent research progress.^{77,113,114} The three isolated isomers were synthesized, separated and characterized as globular (Cys1-Cys3, Cys2-Cys4), ribbon (Cys1-Cys4, Cys2-Cys3) and bead (Cys1-Cys2, Cys3-Cys4) isomer for several α -conotoxins, e.g. α -GI, α -AulB and α -Vc1.1 with distinct pharmacological activities.^{113–115} Furthermore, in 2002 Dutton *et al.* reported for the first time a non-native isomer of α -AulB that exhibits ten times higher biological activity compared to the native connectivity, representing a novel method for optimization of pharmacological properties.¹¹⁵ However, for conotoxins with six cysteines 15 isomers are possible indicating the higher complexity for conopeptides bearing three disulfide linkages such as μ - and δ -conotoxins. The whole problem gets more complicated for an increasing number of cysteines, e.g. for four disulfide bonds 105 isomers are possible.¹⁰⁸ Because of this difficult intricacy and similar properties of the different disulfide isomers information about μ -conotoxins are rare or not existing in the case for δ -conotoxins.^{96,116} In general, different isomers were isolated after oxidative folding and identified via NMR structure elucidation. Determination of the disulfide linkage was performed applying mass spectrometry and MS-sequencing after different approached of sample preparation. However, only two or three isomers could be identified from the two μ -conotoxins μ -KIIIA and μ -PIIIA (Figure 7) and have been described to this day.^{96,116}

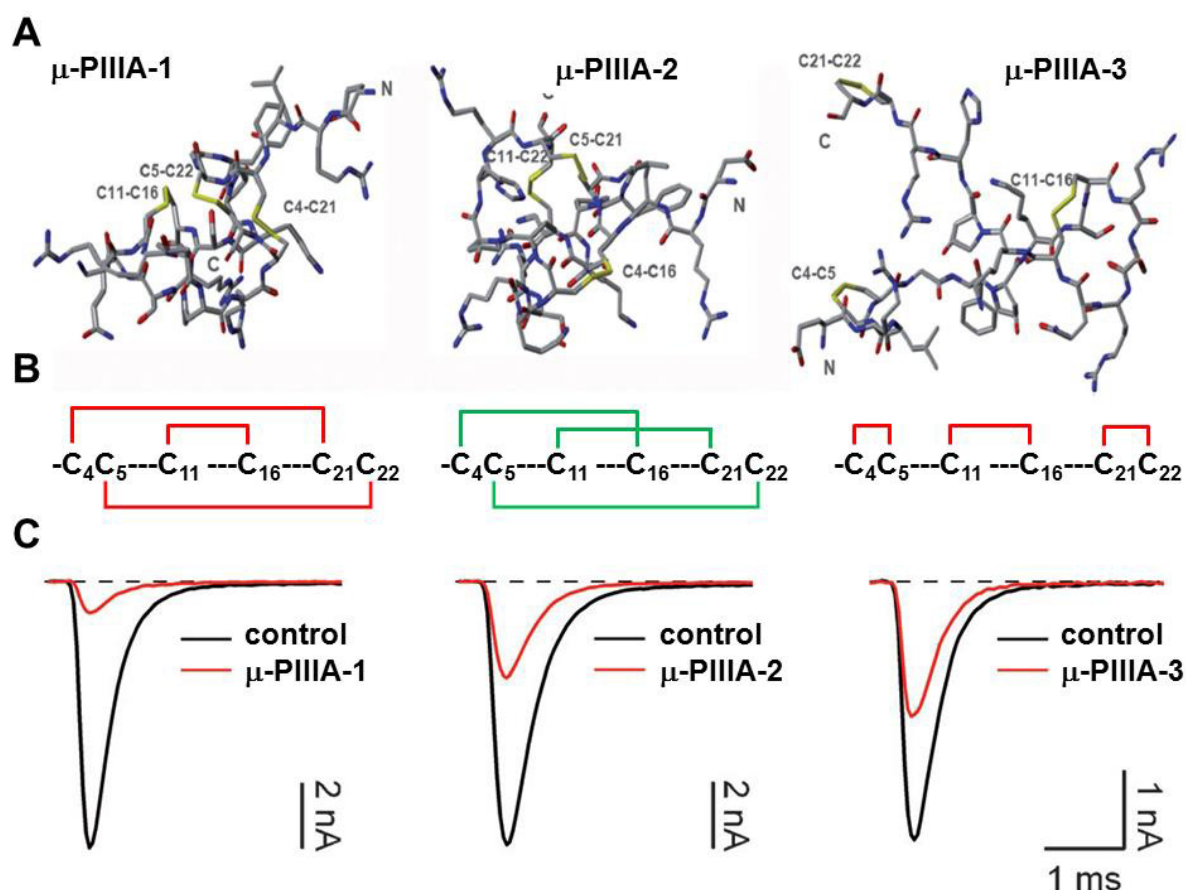


Figure 7. Topology of the three isolated μ -PiIIA isomers from oxidative folding and structures derived from NMR elucidation (A) with the respective disulfide connectivity (B). The individual blocking of VGSC Nav1.4 is described below (C) with the highest bioactivity for a non-native isoform. ^{Modified from 96}

These time consuming and exhausting researches show that not only synthesis but also following structure elucidation is difficult (chapter 2.4). Approaches performed so far only revealed the disulfide linkages observed during *in vitro* studies leaving the question open for *in vivo* conditions in the end.

In addition, the exchange of cysteine with selenocysteine is a further possibility for selective oxidation because of the different redox potential. Extensive research with partially replaced disulfide crosslinks, i.e. with a diselenide bridge, was performed for several peptides such as BPTI, μ -SIIIA and various α -conotoxins indicating the large potential for substituting distinct linkages.^{113,117–119} Beside the previous mentioned chemical synthesis methods for conotoxins also biological methods such as peptide expression in *E. coli*, which was reported for the two conotoxins α -LvIA and LtXVIA from the M-superfamily, show future potential. Nevertheless, both peptides lack of difficult posttranslational modifications hampering conotoxin expression in general.^{120,121}

2.3.1. Protecting group strategies

Protocols for different PG strategies developed and established in the last decades enable regioselective oxidation of highly disulfide linked peptides.^{1,5,28,106} The advantage of this method compared to the following ones is the selective formation of a disulfide bond from two selectively protected cysteine residues after their deprotection and an oxidation procedure. This allows stepwise formation of disulfide connectivity to the intended isomer. However, reaction conditions have to be chosen carefully to avoid breaking or scrambling of the already formed bridging.^{1,5} Furthermore, some PGs can be removed by oxidative cleavage conditions resulting in the disulfide rather than the thiol compound. Some approaches can be performed on-resin but most common procedures are off-resin approaches after cleaving the peptide with partially or fully protected cysteine residues.¹⁰⁶ According to their stability and lability for Fmoc-peptide synthesis the cysteine PG trityl (Trt), acetamidomethyl (Acm), 4-methylbenzyl (Meb)/4-methoxybenzyl (Mob), *tert*-butyl (tBu), and *tert*-butylthio (StBu) are favored.^{5,106} Additionally PGs with special cleavage conditions enable the combination with the common cysteine PG such as enzyme-cleavable (phenylacetamidomethyl, Phacm), photocleavable (coumarin- or nitrobenzyl-derivatives) or safety-catch (4,4-bis(dimethylsulfinyl)benzhydryl, Msbh) PGs.¹⁰⁶ However, their application is also limited for several reasons like side effects or cleavage conditions.^{1,5,28,106}

Research on various α -conotoxins indicates sequence dependencies as well as reaction conditions can affect the yield of the intended isomer.^{114,122,123} This shows the intricacy of disulfide rearrangement even with only two disulfide connectivity. While Akaji *et al.* reported no disulfide shuffling during preparation of α -MI the same PG strategy (Acm and tBu) resulted in ambiguous isomers for the preparation of α -GI performed by Szabo *et al.* revealing the coherence of several parameters like PG order of cysteine residues, reaction time or temperature.^{122,123} Similar to the previous mentioned α -conotoxins, various approaches of selective folding were developed and applied to peptides with six or more cysteine residues like the conotoxins μ -SIIIA,¹²⁴ δ -CnVID,¹²⁵ α -SI dimer¹²⁶ or other disulfide-rich molecules such as hepcidin.^{124,127} Again, in some cases the folding resulted into another isomer, although the reasons for this disulfide rearrangement are still not fully understood and seem to be highly depending on peptide sequence (especially for adjacent cysteine), PG strategy and reaction conditions.^{123,127,128} Finally, elucidation of the disulfide linkage is still a challenging and demanding mission and will be discussed in more detail in chapter 2.4.

2.3.2. Oxidative folding

The so-called oxidative folding approach was established by Creighton to elucidate the folding pathway of BPTI.^{129,130} This approach differs from the conformational folding that is based on unfolding of the native species followed by refolding of the peptide to the native structure. The process is monitored with infrared spectroscopy, NMR, fluorescence spectroscopy or circular dichroism.^{10,109,131,132} On the other hand oxidative folding of proteins allows more detailed information about the folding intermediates, kinetics and folding pathway. This is achieved by folding of reduced, unfolded peptide in buffer solution containing oxidizing agents resulting in the native, restricted structure by formation of the respective disulfide bonds. Hereby the folding intermediates can be isolated and subjected to analytical by HPLC and other methods for characterization and structure determination (chapter 2.4).¹⁰⁹ Different folding pathways are discussed nowadays, whereof three different types incorporate most folding processes from reduced to native state, namely the two extremes BPTI- and hirudin-like folding model or a mixture of both (BPTI/hirudin-like model).¹⁰⁹ Whereas the BPTI folding pathway proceeds exclusively via distinct intermediates bearing the native connectivity, the hirudin folding progresses via various intermediates not only exhibiting the native linkage but non-native isomers depending on the oxidation condition. For this peculiar folding behaviour three condensation models exist depending on the native-like structured parts that form secondary structured elements prior to folding. These elements are anchored in the primary amino acid sequence and determine the accessibility of the thiol groups for the respective native structure. Therefore the framework model represents the BPTI folding with high amount of secondary structure elements, hydrophobic collapse for hirudin-like folding without structural information and nucleation-condensation for the BPTI/hirudin-like folding with structural information.¹⁰⁹ Protein folding with or without cysteine containing peptides and proteins is an important issue because peptide/protein misfolding can result in several severe diseases.^{133,134} Computational *in silico* studies help to understand the folding mechanism in detail using funnel-shaped energy landscapes with different levels from the non-folded structure on top, intermediates in between and the native folded structure at the bottom of the funnel.^{135–138}

However, conotoxins appear to fold in a hirudin-like pathway whereas isolation and characterization of non-native isomers was possible upon oxidative folding as described previously (Figure 7).^{96,109} However, such isomers have also been reported during oxidative folding of several other cysteine-rich peptides such as hirudin,^{109,139} amaranthus α -amylase inhibitor,¹⁴⁰ cardiotoxin III,¹⁴¹ and tridegin.¹⁴² Concluding, these *in vitro* findings are not always consistent with Anfinsen's dogma whereby only one native isomer possessed the bioactive

conformation.^{107,108} Still, *in vivo* conditions can differ because organisms developed several tools to assist folding, e.g. small molecule oxidants or protein disulfide isomerases (PDI).^{1,143} These PDIs constitute a large part of proteins in the venom duct of cone snails indicating the high impact on the folding procedure of conotoxins.¹⁴³

Nevertheless, this common method is often weapon of choice for solution oxidation and conditions were optimized for several conotoxins with three disulfide connectivity including ω -MVIIA,¹⁴⁴ μ -SIIIA,¹⁴⁵ μ -PIIIA⁶³ and δ -PVIA.⁶⁷ Although there are miscellaneous oxidation methods existing, the most common technique is applying redox agent mixtures of reduced and oxidized glutathione (GSH/GSSG) or other oxidizing agents like cystine or cystamin in slightly basic (pH 7.5 - 8.5) buffer solutions and low peptide concentrations to prefer intra- rather than intermolecular reactions.^{1,5} Alternatively, dimethylsulfoxide (DMSO) can be used for oxidation due to the possible application in a pH range of 1 – 8 and it's capability to improve solubility properties especially in the case of hydrophobic peptides. However, side reactions can occur at methionine, tryptophan and tyrosine residues and removal of large amounts of DMSO can complicate purification.^{5,145}

2.3.3. Room temperature Ionic Liquids for peptide oxidation

Ionic liquids (IL) are so called molten salts and room-temperature ionic liquids (RT-IL) are liquid below 100°C.¹⁴⁶ They generally consist of a cation poorly coordinated by an anion, hindering the formation of ordered crystal lattice and resulting in unique physicochemical properties and liquid behaviour.¹⁴⁶ One of the ions is generally a bulky, organic molecule with delocalized charge that is neutralized with a counter-ion.¹⁴⁶ Therefore, a lower melting point compared to common crystalline structures is obtained.¹⁴⁶ As a result RT-ILs are liquid at ambient temperatures. They can be divided in four groups based on the cation namely alkylammonium-, dialkylimidazolium-, phosphonium- and N-alkylpyridinium-based ILs.¹⁴⁷ In the last decade RT-ILs were established for a wide range of applications because of exclusive properties tailored from the containing ions such as negligible vapour pressure, low volatility, high thermal conductivity, large electrochemical window, solvating potential, non-flammable, recyclable and lower toxicity (compared to conventional organic solvents).^{146–148} Therefore, they possess huge potential as environmental friendly reaction media, so called “green solvents” but due to their designable characteristics not all ILs have the quality of green solvents.^{147,149,150} Nevertheless, RT-ILs have been established in chemical synthesis,¹⁴⁹ catalysis,¹⁴⁸ biotransformation,¹⁴⁶ biocatalysis^{151,152} biomass conversion^{152,153} and further applications.^{154,155} The field of peptide synthesis and chemistry as well as protein

2. Theoretical background

interactions (e.g. enzymes) has obtained special interest in this context and RT-ILs are used for different reasons.¹⁵⁶ Protein activity and structure stabilization of enzymes mediated by RT-ILs can be explained with specific ion effects which evolved from the Hofmeister series founded in 1888.^{157–160} Since then chaotropic (weakly hydrated, low charged) and kosmotropic (strongly hydrated, highly charged) ions used for preparation of RT-ILs are described and have stabilising or destabilising effects on enzymes in neat or aqueous RT-ILs, resulting in a change of bioactivity.^{157,161,162}

Another application of RT-ILs as an alternative reaction media is the coupling of amino acids to produce dipeptides in higher purities compared to conventional organic synthesis.¹⁶³ Furthermore, RT-ILs are utilized as resin during SPPS for immobilization of amino acids and coupling procedures to synthesize peptides such as the pentapeptide sansalvamide A or even longer ones with 15 residues.^{164,165} In addition, RT-ILs are used for oxidative folding of disulfide rich peptides without oxidizing agents. In 2009 Miloslavina *et al.* reported the first conotoxin oxidation performed in RT-ILs with the advantage of solving higher concentrations of hydrophilic (μ -SIIIA, μ -PIIIA) and hydrophobic (δ -EVIA, δ -SVIE) peptides in the respective RT-IL (1-ethyl-3-methylimidazolium acetate, [C₂mim][OAc]) and higher yields compared to buffer solutions.¹⁶⁶

RT-ILs are used for several analytical purposes such as NMR, mass spectrometry (MS, especially as matrices), electrochemistry or chromatography.^{156,167–169} For instance, tandem mass spectrometry from peptides out of RT-ILs for Electrospray ionization (ESI)-MS and Matrix-assisted laser desorption ionization (MALDI)-MS was developed.^{170,171} For the latter case the RT-IL was prepared with an alkyne moiety which was able to open and modify the disulfide cysteine by a light induced reaction after dissolving the sample in the RT-IL.¹⁷¹ Afterwards in-source decay (ISD) was directly applied performing fragmentation of the derivatized peptide out of the RT-IL matrix mixture and fostering the formation of fragment-ions compared to classical matrices like 1,5-diaminonaphthalene, 2,5-dihydroxybenzoic acid and α -cyanohydroxycinnamic acid.^{171,172}

2.4 Analytical methods for multiple disulfide-bridged peptides

Besides milestones in peptide and protein synthesis evolved in the past decades analytical methods co-evolved for peptide characterization correspondingly. Therefore, the Nobel prize in chemistry 2002 was awarded to Kurt Wüthrich, John B. Fenn and Koichi Tanaka for their pioneered research in the field of nuclear magnetic resonance (NMR) and mass spectrometry (MS) applied to biological macromolecules.¹⁷³ Recent progress and high

importance of these techniques are related to structure- and disulfide elucidation of disulfide-rich peptides (chapter 2.4.1-3). In general, analytical methods for elucidation of disulfide arrangement can be divided into destructive and non-destructive methods depending on the sample treatment and therefore the possible sample recovery. (Figure 8).¹

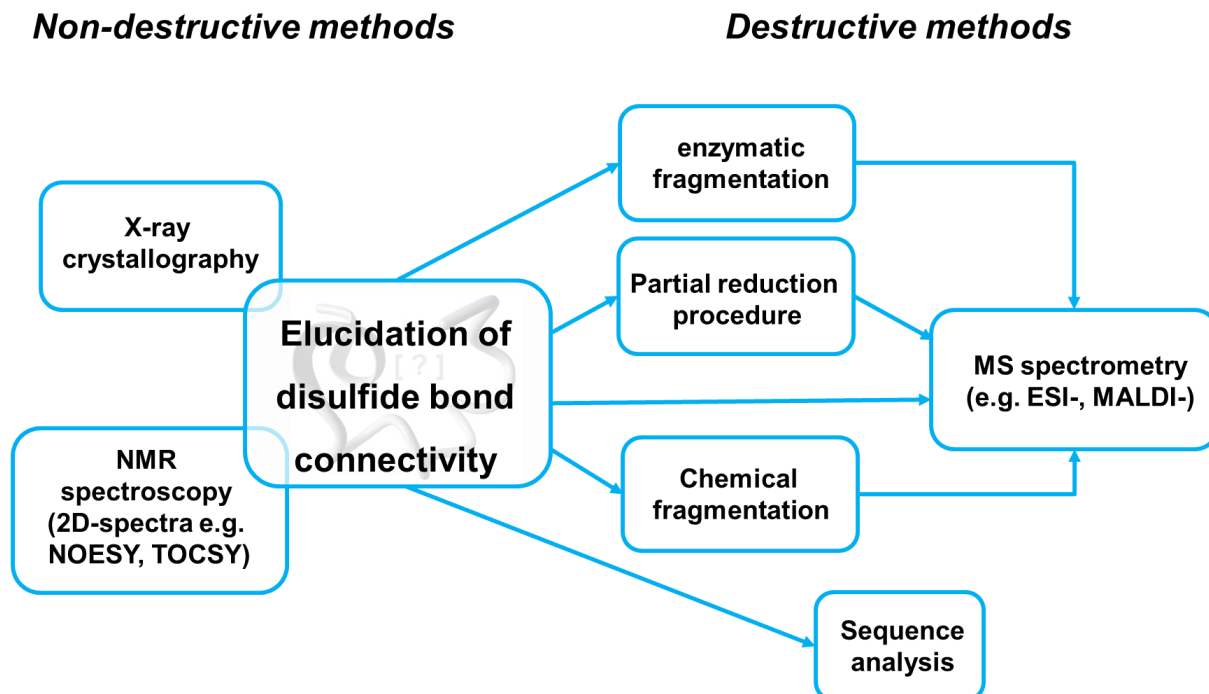


Figure 8. Analytical methods for the elucidation of disulfide bond connectivity of small, disulfide-rich peptides and respective characterization opportunities. Samples subjected to X-ray crystallography or NMR can be characterized by destructive methods afterwards because recovering is possible, which is not possible vice versa. Modified from 1

Both non-destructive approaches (X-ray crystallography, NMR spectroscopy) result in complete 3D-structure models crucial for SAR studies.¹ For X-ray crystallography a major issue is the preparation of suitable crystals from a peptidic or protein solution.¹⁷⁴ Another aspect that has to be considered is that during these measurements the sample is solid and not in solution which can lead to misleading results compared to solution NMR.¹⁷⁵ Solid-state NMR is rarely used for the elucidation of disulfide bridges for the same reasons as mentioned for crystallography. Therefore, solution NMR is the method of choice for small peptides (< 5 kDa) indicated by the structures available, since ~80% of the respective PDB entries are based on this technique. However, determination of the disulfide connectivity is not always possible for each structure and therefore, the final disulfide pattern is unclear or ambiguous isomers are possible for a distinct structure derived from NMR.¹⁷⁶ The sample amount needed for non-destructive methods is typically much higher compared to the destructive techniques. This issue was solved by advances in peptide synthesis to generate sufficient sample amounts that allow sequence identification and structure determination.¹⁷⁷

2. Theoretical background

On the other hand minimal sample amounts especially obtained from native material are analyzed with MS-based techniques nowadays which evolved in the progressive development of different combined “omics” approaches (e.g. venomics, proteomics) representative for the destructive methods.^{14,173,178} In the venomics approach the isolated samples can also be subjected to Edman degradation for identification of the primary sequence and disulfide elucidation.^{179–181} Nowadays, the preferred LC-MS based tandem mass spectrometry combines the chromatographic separation of the crude venom with the benefit of peptide sequencing via tandem mass spectrometry (e.g. LC-ESI-MS/MS).^{1,182} Compared to NMR structure elucidation identification of a single peptide toxin sequence is achieved much faster exploiting the unique advantages of two established ionization methods: ESI- and MALDI-MS or MS/MS (chapters 2.4.1 and 2.4.2). Although both methods are MS-based, the spectra obtained from each method are often unique for each peptide or protein. Interestingly, some can be analyzed with both methods but others are peculiar for the respective method applied using their characteristic advantages and limitations.^{48,178,183,184} In addition, both methods are able to quantify compounds, as was described for anthrax, up to attomol/mL range.¹⁸⁵ Data obtained from MS/MS experiments are tremendously high and can be difficult to evaluate, especially after enzymatic digest and fragmentation.¹⁸⁶ Identification of peptides or proteins is usually performed using programs or databases with respective information available from measured or calculated experiments. Expertise in evaluation of these data is necessary to prevent false data interpretation, which includes time-intensive manual data evaluation if required.^{178,186,187} Nevertheless, algorithms are also existing for the prediction of disulfide patterns of peptides and proteins.¹⁸⁸

For the final elucidation of the disulfide linkage in peptides or proteins by destructive methods, sample preparation is highly essential.^{178,189} Identification of the former disulfide bridge is typically determined after enzymatic digest of the peptide according to enzyme-specific fragments.^{178,189} Comparing reductive and non-reductive conditions can clarify the bridged fragments of the whole peptide with a particular bridge opened or closed.^{142,189} Especially larger peptides and proteins were subjected to this method whereas a trypsin-based digest has the advantage of cleaving behind a positively charged residue which is ideal for detection via ESI-MS which might be hampered otherwise.^{189,190} When suitable enzyme cleavage positions are missing in the peptide sequence, partial reduction together with derivatization of the free thiol function is applied for disulfide elucidation.¹⁹¹ Several protocols have been developed based on different reagents like iodacetamide (IAA), N-ethylmaleimide (NEM) and 4-vinylpyridine that react with the free thiol group of cysteines after reduction with a reducing agent like TCEP or DTT.^{190–193} The elucidation of the derivatized cysteines can be performed by Edman degradation or MS.^{190–192} However,

disulfide scrambling during sample preparation can be challenging and parameters have to be optimized to avoid this, for example by using acidic reaction conditions.¹⁹²

The following chapters will focus on LC-ESI-MS, MALDI-MS and NMR especially for disulfide elucidation including advantages and limitations of the corresponding method.

2.4.1 LC-ESI mass spectrometry

Electrospray ionization (ESI) is one of the two destructive methods established for analytics of biomolecules.^{194,195} The major challenge is the soft ionization of the large biomolecules and the transfer to the gaseous phase.^{178,194} This was realized for solved biomolecules by an effect during evaporation of small droplets bearing a high positive charge. This leads to a disruption into smaller charged droplets and in the end to charged molecules only without solvent.^{178,194} Proteins and large peptides are usually detected as highly charged ions.^{173,194} The great benefit of this technique is the direct combination of liquid chromatographic sample purification and MS for identification of each single fraction.¹⁹⁶ Sequencing of the respective peptide, protein is typically performed with collision-induced decay (CID) in which a collision cell is used for fragmentation and isolation of the peptide ions.¹⁹⁷ Two types of CID can be differentiate depending on the energy applied namely low energy CID and high energy CID (typically in ion trap MS) with a shorter activation time compared to conventional.¹⁹⁸ Today, several devices utilizing ESI were used with different instrumental configurations e.g. quadrupole time-of-flight (Q-TOF), ion trap and Fourier transform ion cyclotron resonance mass spectrometer (FT-MS) each providing unique advantages and limitations.¹⁷⁸

In general, the use trypsin digest represents another advantage because of the cleavage site resulting in positively charged fragments that are easier to detect compared to other fragments observed by alternative enzymes.¹⁹⁵ Although the disulfide linkage of μ -KIIIA was elucidated with tryptic digest in combination with CID-MS/MS for many other conotoxins this approach fails because they lack of suitable cleavage sites or ambiguous fragments obtained.^{5,116} More often a partial reduction and derivatization protocol is required prior to MS/MS sequencing as mentioned previously. The elucidation of α -Si by Yen *et al.* was performed by partial reduction with TCEP in combination with alkylation using NEM and afterwards complete reduction and derivatization with IAA.¹⁹² The distinct alkylation of the thiol-groups revealed the disulfide connectivity of the examined fraction. It has to be mentioned that a previous approach using maleimide-biotin instead of NEM for derivatization failed because separation of the products was not possible.¹⁹² Nevertheless, the disulfide crosslinking of several conotoxins could be determined with this method after various preparation protocols (Table 2). But careful considerations are required because it is

2. Theoretical background

reported that during preparation disulfide scrambling can occur resulting in unexpected fragments and a false determination of the disulfide bridging in the end.^{192,199} Therefore, different approaches are often used together for validation and reproducibility of the obtained fragments.^{184,199}

2.4.2 MALDI mass spectrometry

Matrix-assisted laser desorption ionization (MALDI) represents the second MS method established for soft ionization of biomolecules.²⁰⁰ In this case the macromolecule is co-crystallized with a distinct matrix.²⁰⁴ Ionization is performed afterwards with a laser suitable for the applied matrix, which evaporates it together with the intact biomolecule. With this approach higher masses can be investigated compared to ESI-MS and in addition the ionization method is softer compared to ESI-MS.²⁰⁴ One advantage of this approach is the robustness against contamination or sample impurities.^{173,200} Obtained MALDI-spectra are often less complex compared to ESI-spectra because they typically contain singly and doubly charged species, facilitating the evaluation.¹⁹⁰ In general, two techniques for tandem mass spectrometry can be applied during MALDI namely in-source-decay (ISD) or the LIFT-TOF/TOF-MS.^{201,202} The latter has the advantage of preselection of a precursor mass prior to fragmentation, whereas ISD reveals unique fragmentation pathways able to break disulfide bonds during fragmentation.^{201,202} Changing the matrix is a further possibility to increase the information obtained from a complex sample like crude venom.²⁰³ Application of 1,5-diaminonaphthalene (1,5-DAN) together with α -cyano-4-hydroxycinnamic acid (HCCA) as matrices reveals the amount of cysteine bridges of single samples in the venom mixture by comparison of reduced and oxidized species because of the reducing properties of 1,5-DAN in contrast to conventional matrices.²⁰³ But usually previously mentioned sample preparation is also necessary to digest or reduce and alkylate the peptide for disulfide elucidation. It should be mentioned that during the laser ionization process gas-phase scrambling of disulfide bridges can occur.²⁰⁴ Nevertheless, MALDI-MS/MS raised during the last decades because of new approaches established including technical improvements as ISD and Lift. Combination of MALDI- and ESI-MS/MS was used for characterization of the bromotryptophan-containing conotoxin Mo1274 revealing the disulfide connectivity Cys1-Cys3 and Cys2-Cys4.²⁰⁵ Comparison of the results obtained from MALDI- and ESI-MS for native material from *C. purpurascens* a higher number of masses was detected in the ESI-approach.⁴⁸ This was explained by poor ionization of the peptides examined.⁴⁸ MALDI-MS/MS holds also great perspectives besides disulfide elucidation. For instance, one special application for MALDI-MS/MS is imaging of tissue layers.²⁰⁶ With this method complex

2. Theoretical background

molecular structures on surfaces can reveal detailed information of one distinct or several indicative masses on a certain area after covering with matrix.²⁰⁶

Conotoxin	Disulfide bridges	Sample preparation	Analysis method
α -SI ¹⁹²	2	partial reduction and differential alkylation	ESI-MS/MS
α -CnIa, α -GI ¹⁹⁹	2	partial reduction and differential alkylation or oxidation to cysteic acid	ion mobility MS/MS CID
Mo1274 ²⁰⁵	2	partial reduction and differential alkylation	ESI-FT-ICR MS and MS/MS
α -GI ¹¹⁴	2	intact peptide	NMR
Ar1232, Ar1248, Ar1311, Vi1360, Vi1358, Ar1446, Ar1430 ²⁰⁷	2-3	intact peptide and tryptic digest	ESI-MS/MS, CID for disulfide bond cleavage
μ -GIIIA, μ -GIIIB, μ -GIIC ²⁰⁸	3	partial reduction, carbamido-methylation and different digests	Edman degradation
μ -KIIIA ¹¹⁶	3	tryptic digest, intact peptide	LC-ion trap MS/MS CID, NMR
μ -SIIIA ^{91,95}	3	intact peptide	NMR
μ -PIIA ²⁰⁹	3	partial reduction and differential alkylation	Edman degradation
μ -PIIA ⁹⁶	3	tryptic digest, intact peptide	MALDI-MS/MS, NMR
μ -PIIA ⁶⁴	3	intact peptide	NMR
δ -EVIA ²¹⁰	3	partial reduction and differential alkylation, intact peptide	Edman degradation, NMR
δ -GmVIA ²¹¹	3	partial reduction and differential alkylation	Edman degradation
BtIIIA ²¹²	3	partial reduction and differential alkylation, intact peptide	Edman degradation, NMR
MrIIIA, TxIIIC ⁶⁵	3	partial reduction and differential alkylation	Edman degradation
Mo3964 ²¹³	3	partial reduction and cyanilation, alkaline cleavage, intact peptide	MALDI-MS, NMR

Table 2. Comparison of disulfide elucidation methods and sample preparation techniques applied to respective conotoxins.

2.4.3 Two-dimensional NMR spectroscopy

Nearly 60 years ago Saunders *et al.* reported the first NMR experiments with a biological macromolecule, Ribonuclease.²¹⁴ Since then technical progress in superconducting magnets and progressive computational power for instrumental control and processing of spectra led from simple 1D to 3D NMR techniques.²¹⁵ This includes different homonuclear (¹H, protons) and heteronuclear (e.g. ¹³C, ¹⁵N) couplings with distinct structural information. Therefore isotope-labelling is often extended to proteins especially for structural or kinetic investigation of intra- and intermolecular interactions.²¹⁵ Furthermore, NMR elucidation became extremely important for structure-based drug design not only for small molecules but also for larger peptides.^{175,216} Compared to common through-bound approaches such as correlation spectroscopy (COSY) or Total correlation spectroscopy (TOCSY) that give limited information about protons in close proximity (spin-spin coupling), the development of Nuclear Overhauser Enhancement Spectroscopy (NOESY) facilitates conformational observations also for protons correlating through-space (dipole-dipole coupling).^{175,216} Besides the already mentioned experiments several other correlation methods were established like rotating frame Overhauser effect spectroscopy (ROESY) that enhances limitations and enables high-resolution NMR measurements.²¹⁷ However, evaluation of the recorded spectra is still time consuming and demanding because combination of different experiments is taken into consideration, because each provides peculiar information for structure analysis. The known peptide sequence resonance assignment is performed together with a sequential assignment strategy based on the unique chemical shift of the included amino acids and their characteristic cross peaks in the spectra. The complete assignment provides information about distance constraints and dihedral angles between residues eventually resulting in a distinct structure.^{175,218}

Nowadays, for the elucidation and calculation of structures derived from NMR spectra various software programs are established and facilitate procedures such as CCPN or YASARA.^{219–221} Special interests within the field of structure elucidation is the visualization of complex macromolecular structures.^{222,223} This gives the possibility of representing a plethora of information to a broader community of life scientists.²²² Furthermore, NMR gives the opportunity to observe peptide or protein folding with atomic resolution to elucidate the folding mechanism.¹¹²

Within NMR structure determination the elucidation of disulfide connectivity is even more sophisticated but essential because at least one disulfide bond is present at 23% of the ~62000 protein structures in PDB.^{224,225} For conopeptides, increasing numbers of several thousands of toxin primary sequences observed from high-throughput-screenings have been

described, whereas only a fractional amount of NMR-structures (164 from conoserver, 20.12.2016) are available.⁶⁸ This indicates the limitations in structure determination because of high expenditure of time and labour for generating such structures. Obviously the structures are also crucial for SAR-studies.⁴

Structure elucidation of several conotoxins elucidated different structures for the respective peptides. Different isomers could be clearly separated and characterized via NMR by Gehrmann *et al.* for α -GI and by Carstens *et al.* for various other α -conotoxin.¹¹⁴ For more complex μ -conotoxins structure elucidation is also available for several peptides (Table 2). Although they prefer a distinct disulfide connectivity with two adjacent cysteines, for other M-conotoxins also other disulfide crosslinking could be identified as well, e.g. BtIIIA, MrIIIA or TxIIIA within.^{212,226,227} All venomous animals produce a large variety of peptidic macromolecules from which ~70% bear two or more disulfide bridges representing a common feature in nature.^{228,229} Due to the high structural diversity depending on the primary sequence, including adjacent cysteines, the number of cysteines together with the formation of isomers and the absence of a sulphur isotope with beneficial NMR properties the unequivocal structure elucidation is hampered. The two major issues for elucidation are overlapping chemical shifts during NOESY of bridged cysteines and, on the other hand, adjacent cysteines, i.e. the connected cysteine is in close proximity to another not linked cysteine residue.²²⁸ Trusting only one technique can result in misleading structures as was reported for the NMR structure of hepcidin derived from NOESY.^{127,228} Therefore, several *in silico* methods developed for the prediction of the disulfide connectivity based on the existing data should be applied to overcome the previous mentioned problems. However, these computational programs focus their prediction algorithm on available structures, which might be revisited like mentioned for hepcidin, μ -KIIIA and neopetrosiamides.^{116,127,228,230} Nevertheless, especially for peptides with three or more disulfide bridges different isomers still lack standards, including all possible isomers for a preserved sequence.^{176,231} One possibility of structure determination regarding the disulfide arrangement is isotope labelling with $^{15}\text{N}/^{13}\text{C}$ cysteine residues and 2D versions of standard triple resonance ($^1\text{H}, ^{13}\text{C}, ^{15}\text{N}$) NMR experiments and 2D [$^{13}\text{C}, ^1\text{H}$] HSQC, which might also be sophisticated in the case of two or more adjacent cysteines.²³²

3. Thesis outline

Conotoxins may represent an interesting research target for the treatment of pain due to their impact on ion channels and/or receptors. This work aimed at the investigation of the folding behaviour of disulfide-bridged conotoxins under various oxidation conditions, a comprehensive analytical characterization of the resulting products and on their structure-activity relationship. In order to pursue this goal, several approaches should be examined for the preparation of these target molecules, namely oxidation in conventional buffer systems, oxidation in Ionic Liquids, and selective cysteine bond formation by successive oxidation using an orthogonal protecting group strategy. The individual products should then be characterized by various analytical techniques such as HPLC, NMR, LC-ESI-MS/MS and/or MALDI-MS/MS. Furthermore, the impact of the disulfide-connectivity of the respective conotoxin on the biological activity at VGSCs should be considered as well. In addition, several computational methods should be applied to elucidate the folding behaviour of the peptides as well as the interaction of a respective conotoxin and selected VGSCs.

As a starting point, peptide oxidation in room-temperature Ionic Liquids and the analytical characterization of the resulting products are to be investigated.¹⁶⁶ Therefore, the influence of different IL compositions as well as educt concentration, temperature, and water content on product formation of different conotoxins should be analyzed. Validation and limitations of distinct methods (e.g. MS/MS) suitable for peptide chemical and structural analysis are in the scope of the second main topic of this thesis aiming at a direct comparison of these techniques. Here, μ -PIIIA containing three disulfide bonds should be utilized as model peptide for the selective synthesis of the 15 possible disulfide isomers, which are to be compared to the three isomers earlier isolated from a self-folding approach.⁹⁶ Thus, selective formation of disulfide bridges should be performed using an orthogonal protecting group strategy and successive disulfide bond formation. In addition, the impact of the disulfide connectivity on conformation and, in turn, binding to and inhibition of the skeletal ion channel Na_v1.4 should be examined. The obtained information should be combined and evaluated to derive structure-activity-relationships. In the end, the question whether one method alone can distinguish between all isomers or not should be answered. This is crucial to verify the results from the characterization of peptides containing multiple disulfide bonds in order to finally be able to establish standardization for the analysis of such disulfide isomers. This work demonstrates the essential need for the generation and application of standards and standard methods to provide unequivocal evidence of such complex molecules.

The present thesis is subdivided in six parts according to the publications included herein.

Part I (chapter 4.1) explains several aspects of the application of ILs in peptide chemistry. In this review the advantages and disadvantages of ILs as novel reaction media are explained considering reports of the last decades with respect to synthesis, purification and analytical characterization. Current developments and perspectives are discussed in detail within this topic.

Part II (chapter 4.2) presents the results of the two reactions namely oxidative folding and native chemical ligation of cysteine-containing peptides in ILs. Furthermore, the impact of reaction conditions such as temperature and water content on product formation are reported.

Part III (chapter 4.3) encompasses the results of the application of various ILs on the oxidation of distinct model conotoxins containing two or three disulfide bonds. In addition, different IL mixtures and reaction conditions are analyzed using μ -SIIIA as model peptide. In this case, disulfide bonds were elucidated via LC-ESI-MS/MS after partial reduction and alkylation to confirm the native disulfide connectivity comparing the buffer- and IL-oxidized peptide. The impact of IL ions adhering to the peptide was demonstrated by differences regarding the biological activity of μ -SIIIA at VGSC Na_V1.4 and yet, as a substrate for proteolytic degradation by the enzyme trypsin.

Part IV (chapter 4.4) contains the characterization of the essential binding site of δ -EVIA at VGSCs and different chimeras of Na_V1.4 and Na_V1.7. Prior to the experimentally performed bioactivity tests and the molecular modelling studies, the structure of δ -EVIA was confirmed with solution NMR and LC-ESI-MS/MS.

Part V (chapter 4.5) presents the results of the folding behaviour of a set of μ -conotoxins examined using molecular dynamics simulation of the NMR structures and also by computer visualistics studies with the aim to elucidate the influence of individual amino acids in the sequence to the folding process. Especially, the influence of individual disulfide bonds on the structural stability of the peptides after stepwise exclusion of the respective bridge was in the focus. The partially unfolded species with one disulfide bond removed show similar stability to the folded conotoxin which might indicate a possible retention of the biological activity. Although the different disulfide bridges increase the conformational stability of all structures, their positions within the molecule frequently change keeping the basic residues responsible for blocking of VGSC in a particular orientation on the surface of the peptide structure.

3. Thesis outline

Part VI (chapter 4.6) comprises the results from selective synthesis, structure elucidation and biological characterization of the 15 possible isomers of μ -PIIIA. Comparison of the results obtained with the analytical techniques HPLC, NMR and MALDI-MS/MS should reveal if one method alone is sufficient to distinguish between all isomers. For the first time, structure-activity relationship studies of these isomers are described regarding their inhibitory effect on the ion channel $\text{Na}_v1.4$.

4. Manuscripts

4.1 Ionic Liquid applications in peptide chemistry: synthesis, purification and analytical characterization processes

Review

Authors

Alesia A. Tietze, Pascal Heimer, Annegret Stark, Diana Imhof

Molecules 2012, 17(4), 4158-4185.

doi: 10.3390/molecules17044158

Copyright © 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license .

Preface

Ionic Liquids represent alternative solvents for chemical and biological applications because of their unique properties.^{151,152,233} Ionic Liquids were established in peptide chemistry as novel solvents with improved solubility.^{156,163} Furthermore, purification and analytical procedures represent further applications for ILs with respect to peptides.¹⁵⁶ This article contains an overview about current research and progress in the field of ILs and their broad application with respect to peptide synthesis, purification and analytics.

Review

Ionic Liquid Applications in Peptide Chemistry: Synthesis, Purification and Analytical Characterization Processes

Alesia A. Tietze ^{1,*}, Pascal Heimer ¹, Annegret Stark ² and Diana Imhof ¹

¹ Pharmaceutical Chemistry I, Institute of Pharmacy, University of Bonn, Brühler Straße 7, Bonn D-53119, Germany; E-Mails: pheimer@uni-bonn.de (P.H.); dimhof@uni-bonn.de (D.I.)

² Institute of Technical Chemistry, University of Leipzig, Linnéstraße 3, Leipzig D-04107, Germany; E-Mail: annegret.stark@uni-leipzig.de

* Author to whom correspondence should be addressed; E-Mail: alesia.tietze@googlemail.com; Tel.: +49-228-73-60258; Fax: +49-228-73-60223.

Received: 17 February 2012; in revised form: 22 March 2012 / Accepted: 28 March 2012 /

Published: 5 April 2012

Abstract: This review aims to provide a comprehensive overview of the recent advances made in the field of ionic liquids in peptide chemistry and peptide analytics.

Keywords: ionic liquids; peptide synthesis; peptide analysis; reaction media

Abbreviations:

[BF ₄] [−] = tetrafluoroborate	[C ₁ C ₁ IM] ⁺ = 1,3-dimethylimidazolium
[C ₁ mim] ⁺ = 1-methyl-3-methylimidazolium	[C ₂ mim] ⁺ = 1-ethyl-3-methylimidazolium
[C ₄ C ₁ pyrr] ⁺ = 1-butyl-1-methylpyrrolidinium	[C ₄ mim] ⁺ = 1-butyl-3-methylimidazolium
[C ₆ mim] ⁺ = 1-hexyl-3-methylimidazolium	[C ₈ mim] ⁺ = 1-octyl-3-methylimidazolium
[C _{i,j,k,l} N] ⁺ = tetraalkylammonium	[DEP] [−] = diethyl phosphate
ESI = electrospray ionization	[Et ₃ NH] ⁺ = triethylammonium
[EtOSO ₃] [−] = ethylsulfate	[guan] ⁺ = guanidinium
[Me ₂ PO ₄] [−] = dimethylphosphate	[MeOSO ₃] [−] = methylsulfate
[MOEMIm] ⁺ = 3-(2-methoxyethyl)-1-methyl-imidazolium	
[N(CN) ₂] [−] = dicyanamide	[OAc] [−] = acetate
[OTf] [−] = triflate	[OTs] [−] = tosylate
[PF ₆] [−] = hexafluorophosphate	[SbF ₆] [−] = hexafluoroantimonate
[SCN] [−] = thiocyanate	[NTf ₂] [−] = bis(trifluoromethanesulfonyl)imide
BOP = benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate	

CCA = α -cyano-4-hydroxycinnamic acid	DCC = N,N'-dicyclohexylcarbodiimide
DHB = 2,5-dihydroxybenzoic acid	DIEA = N,N-diisopropylethylamine, or Hünig's base
DMAP = 4-dimethylaminopyridine	DMED = N,N-dimethyl-ethylenediamine
GSH = glutathione	
GTHAP = [1,1,3,3,-tetramethylguanidinium-2,4,6,-trihydroxyacetophenone]	
H ₂ O ₂ = hydrogen peroxide	H ₃ PO ₄ = phosphoric acid
HATU = O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate	
HPLC = high-performance liquid chromatography (high-pressure liquid chromatography)	
IL = ionic liquid	ILM = ionic liquid matrices
kDa = kilodalton	LSER = linear solvation free energy relationship
MPG = 2-methyl-2-(<i>p</i> -tolyl)-glycine	MALDI = matrix-assisted laser desorption/ionization
MS = mass spectrometry	MOPS = 3-[<i>N</i> -morpholino]propanesulfonic acid
NCL = native chemical ligation	NMR = nuclear magnetic resonance
OGp = chymotrypsin and trypsin	RP = reversed phase
PyBOP = (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate)	
SCIL = surface-confined ionic liquids	SCm = V8 protease
SPPS = solid phase peptide synthesis	TFA = trifluoroacetic acid
THF = tetrahydrofuran	TLC = thin layer chromatography
TOF = time-of-flight	TRH = thyroliberin (thyrotropin-releasing hormone)

1. Introduction

During the last decade the interest in establishing alternative solvents for organic synthesis to provide safe and environmentally clean processes has increased dramatically. Alternative solvent media that can provide suitable or even better reaction conditions for organic synthesis and which are less hazardous than conventional organic solvents are “molten salts” or so-called ionic liquids (ILs). Numerous reports have been published describing the unique properties of ILs such as negligible vapor pressure, low melting point, special phase behavior, electric conductivity, low flammability, and high thermal stability. The applications range from transition metal-mediated catalysis, e.g., hydrogenation, oxidation, or hydroformylation, to organic reactions such as the Diels-Alder reaction, the Friedel-Crafts reaction, cycloaddition, and esterification [1–3]. In addition, ILs became very attractive in a number of fields involving biological macromolecules, e.g., enzymes and proteins, where they lead to structure stabilization [4] and are applied to biotransformations by enzymatic catalysis [5], protein crystallization [6], and fluorescence quenching immunoassays [7]. In other applications, ILs show advantages due to their favorable solvation properties for polar substrates, including carbohydrates [8,9], proteins [10], and reduced emissions, but also as solvents in analytical chemistry and in chemosensing strategies [11].

In comparison to conventional solvents an IL consists of an anion and a cation, *i.e.*, it possesses a dual character. For this reason, ILs are salts that can provide a suitable medium for biomolecules. Depending on the ionic liquid composition, the solvation environment, they can be adjusted to fit a specific solute and process requirements, such as solubility or phase separation from an organic or an

aqueous phase [12]. By manipulating the ionic solvent properties it is possible to design ILs for special reaction conditions, which in case of enzymatic reactions affect, for example, reaction rate, enzyme selectivity, and substrate solubility [5]. Lately, several reports about the interactions between an ionic liquid and various biomolecules were published. However, the nature and the forces of IL-solute interactions are still too little understood, to answer the questions why ILs stabilize the active form of enzymes/proteins [8,13], or why they provide an ideal platform for the folding of cysteine-rich peptides [14]. In the coming years, these will become more and more attractive in the field of biochemical research.

In 2008 [15], Plaquevent *et al.* presented an excellent overview of IL applications in amino acid and, in part, peptide chemistry, as well as of the use of IL components to produce novel amino acid-based ILs [16,17]. Their intention was to answer the following quite comprehensive questions: Is it possible that ILs can provide a better reaction medium than traditional organic solvents and in which fields of chemistry and biochemistry can they be applied? However, there are still no satisfying answers and many more questions are open due to the large variety of applications offered in this field of research. During the last three years, interest in the use of ILs in peptide and protein research increased immensely. Finding new ways to improve coupling reactions, solubility of peptides and proteins in aqueous/buffer solutions, and their chemical characterization occupied the mind of peptide, protein and analytical chemists. Hence, detailed studies were carried out, including investigations on the influence of peptide/protein concentration on the outcome of a biochemical reaction, the effect of the ionic liquid on the peptide/protein structure, the effect on the function of a peptide/protein of interest, and more generally, on how to increase the stability and reactivity of such biomolecules in ILs.

The main challenge is to select the best IL for a particular process and/or application. This review intends to support advances made in the field of IL applications in peptide chemistry by summarizing the current state in the field and condense the information available in the literature. Indeed, one may ask the question, why we are focusing our attention on peptides? Polypeptides and proteins are present in all living cells in a rather large number and variety. Model peptides are usually used to understand functions of large proteins [18]. However, it is also generally accepted that the effects and behavior of peptides in various reaction media differ from those of proteins. Proteins usually fold to form particular three-dimensional shapes, which determine their actions, while polypeptides are structurally less constrained, *i.e.*, they can adopt many conformations in solution. Thus, proteins possess a rather well-defined tertiary structure that is in many cases missing in the case of peptides. Therefore, we intend to give an overview of the new fields of investigations that appeared during the last years with a focus on the use of ionic liquids in peptide chemistry, including purification, and analysis. Also, this survey represents new trends, which have appeared in the literature following the review of Plaquevent *et al.* [15] four years ago, to help to choose a suitable IL for the desired reactions and characterization techniques.

2. Ionic Liquids in Peptide Chemistry

In this part, we describe the applications of ILs in peptide chemistry: peptide synthesis and modifications of special peptides. We included different synthetic strategies of conventional methods, where organic solvents are used primarily, but focused our attention on ILs as reaction media and the advantages and disadvantages of their use in peptide synthesis. Our report highlights the achievements made in peptide synthesis since the publication of the review mentioned earlier [15].

The choice of the ionic liquid in a peptide-chemical reaction depends on various factors, but most importantly is dependent on: (a) the solubility of the peptide/protein and/or its precursors and (b) the melting point of the ionic liquid. The latter aspect is important if the reaction is to be conducted in the absence of any co-solvent such as water, which would reduce the viscosity of the reaction mixture, but also affect the peptide-ionic liquid interactions. In analogy to other biopolymers, such as cellulose or silk, ionic liquids with rather nucleophilic anions dissolve proteins due to their ability to break the inter- and intramolecular hydrogen bonding network. Similarly, ionic liquids reduce the ability to form peptide-peptide hydrogen bonds, thus lowering the risk of protein denaturation through misfolding and aggregation. However, due to this property, protein-dissolving ionic liquids often feature melting points well above room temperature, which reduces their potential for applications with thermolabile peptides and proteins. While the anion is involved directly in the dissolution process, the cation choice can be used to modulate the properties of the ionic liquid with regards to viscosity, melting point and therefore mass transfer limitations somewhat. Dialkylimidazolium-based ionic liquids belong to the best investigated group of ionic liquids. However, it needs to be noted that questions of toxicity of ILs arise, especially if the peptide is to be used as pharmaceutical. For example, it can be expected that thiocyanide-, methylsulfate or biscyanamide-based ionic liquids will be toxic, while chlorides, sulfates and acetates should be less or non-toxic. In fact, [C₂mim][OAc] is non-toxic and has been proven to be the IL of choice for the dissolution and conversion of the linear, rather hydrophobic peptides [14].

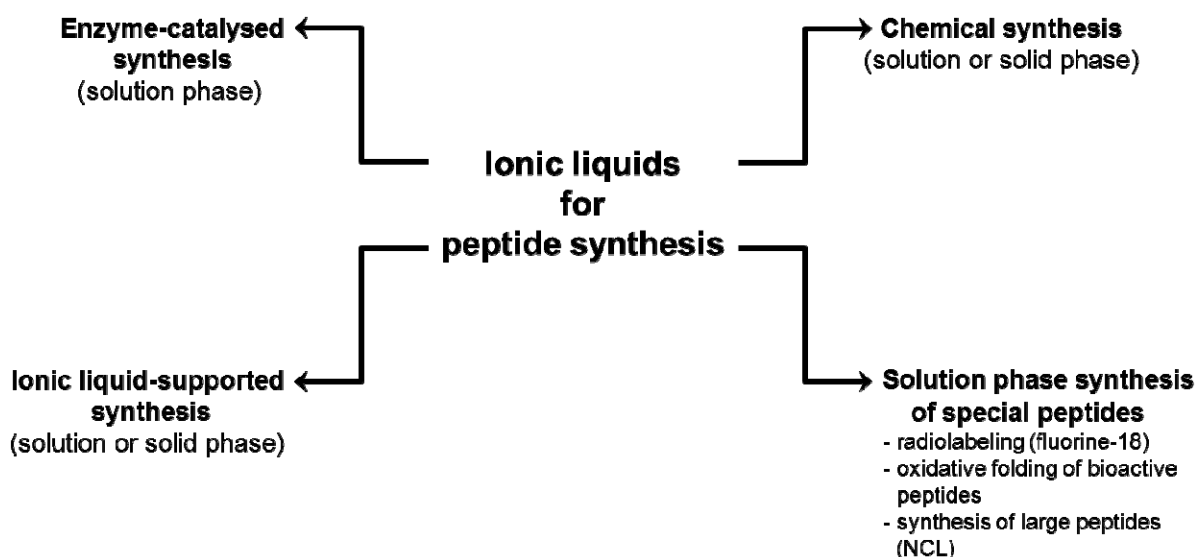
2.1. Peptide Assembly

Nowadays, a large variety of methods is available to produce or obtain peptides, including isolation from natural sources (e.g., extraction from plants, animals and other organisms), recombinant synthesis, and production by chemical or enzymatic synthesis [19–23]. Only a few manuscripts have reported that ILs can be applied in this field of peptide assembly as pure solvents, as co-solvents in aqueous systems or as biphasic systems [8,24], in both, solid and solution phase synthesis as well as in enzyme-catalyzed peptide synthesis (Figure 1).

In general, the choice of the synthetic method is usually determined by the length and the quantity of the peptide of interest. For example, enzymatic peptide synthesis is applied for the production of small peptides, e.g., di- or tripeptides [25], to obtain large quantities in industrial production scale. This method benefits from the high specificity of the reaction, the avoidance of side chain protection of the substrates and the racemate-free formation of the products. Chemical synthetic protocols have focused firstly, on the synthesis of longer peptides by Merrifield solid phase peptide synthesis (SPPS) [26], which is limited to a length of approximately 50 amino acids (though longer peptides can be prepared depending on the primary amino acid sequence), primarily because of the steric hindrance and

accumulation of undesired and truncated sequences during chain prolongation. Secondly, for the synthesis of polypeptides (>50 amino acids) or small proteins the method of segment ligation has been introduced as an alternative for recombinant production (see Section 2.2.3). The combination of chemical and enzymatic synthesis has been described by Guzman *et al.* representing a method that might be quite suitable for selected peptides [23]. In other procedures, the use of a soluble polymer support, which is called “liquid-phase” peptide synthesis, has been demonstrated [27]. This method was successfully applied for the synthesis of several oligopeptides, but the limiting factor here usually is the solubility during peptide assembly and work-up, in particular in solvents such as water or diethylether. In addition, “fluorous-phase” synthesis was established, where fluorous supports (*i.e.*, trialkoxybenzhydryl-type, Wang-type and *tert*-butyl-type support) and fluorous chemistry were used for peptide synthesis exemplified for bioactive peptides, *e.g.*, TRH [28,29]. However, the most common procedure of peptide synthesis still remains SPPS.

Figure 1. Schematic representation of ionic liquid-applications in peptide synthesis.



However, the methodologies applied nowadays clearly suffer from limitations covering solubility, low selectivity and yield, reactivity and steric issues as well as rather high costs of reagents. Thus, there is a high demand for new and/or optimized methods in this field. Indeed, ionic liquids have found their applications in the field of peptide chemistry in the last decade (Table 1). Since it was noticed that ILs improve the solubility of different reactants, separation of the products and the recovery of the ILs was possible [8–15]. ILs became more and more popular in peptide chemistry as summarized in the following sections.

Table 1. Summary of the ionic liquids used in peptide chemistry.

Ionic liquid	Abbreviation	Application	Ref.
1-methoxyethyl-3-methyl-imidazolium hexafluorophosphate or tetrafluoroborate	[MOEMIM][PF ₆] [MOEMIM][BF ₄]	Enzymatic peptide synthesis	[30]
1-butyl-3-methylimidazolium hexafluorophosphate	[C ₄ mim][PF ₆]	Chemical peptide synthesis	[15, 31–34]

Table 1. Cont.

Ionic liquid	Abbreviation	Application	Ref.
1-hydroxyethyl-1-methyl-imidazolium tetrafluoroborate 3-(2-hydroxyethyl)-1-methyl-imidazolium tetrafluoroborate	-	IL supported peptide synthesis	[35,36]
1-butyl-3-methylimidazolium X	[C ₄ mim][X] (X = BF ₄ , PF ₆ , SbF ₆ , OTf, NTf ₂)	Radiolabeling of peptides	[37,38]
1-ethyl-3-methylimidazolium X 1-butyl-3-methylimidazolium acetate	[C ₂ mim][X], (X = OAc, DEP, OTs, N(CN) ₂) [C ₄ mim][OAc]	Oxidative folding and native chemical ligation of cysteine-containing peptides	[14,39]

2.1.1. Enzymatic Peptide Synthesis

Some reports describe ILs as suitable media to increase the solubility, stability and activity of enzymes.[10,13,40–43] For example, Constantinescu *et al.* [41] studied the denaturation of enzymes in ILs including the cations [C₂mim]⁺, [C₄mim]⁺, [C₆mim]⁺, [C₄C₁pyrr]⁺, [C_{i,j,k,l}N]⁺, [guan]⁺ and the anions [SCN][−], [MeOSO₃][−], [EtOSO₃][−], [OTf][−], [NTf₂][−], [N(CN)₂][−]. It was found that the stability at higher temperature (70 °C) was increased compared to aqueous buffer solution, with [C_{i,j,k,l}N]⁺ and [NTf₂][−] showing the best results. These studies were later extended by others, for example to the use of [C₂mim][FSI] (FSI = bis(fluorosulfonyl)imide) for the α-chymotrypsin-catalyzed reaction of Ac-Trp-OEt with Gly-Gly-NH₂ to yield Ac-Trp-Gly-Gly-NH₂ (see also below) [42]. Here, a 16-fold increase in the initial rates for peptide synthesis was observed compared to the reaction carried out in acetonitrile. In addition to these findings, they also found a 17-fold higher activity for α-chymotrypsin in ILs compared to organic solvents.

Besides, other applications of ILs have been reported with respect to proteins, e.g., Summers and Flowers have described the renaturation of proteins in the IL ethylammonium nitrate as a one-step process [44]. This IL prevented aggregation and caused the refolding of proteins, e.g., lysozyme, resulting in its active state with up to 90% of active protein [44]. ILs as suppressors of protein aggregation and useful supports for refolding processes of denatured proteins were also described by Buchfink *et al.* in 2010 [45]. A series of [C₂mim]⁺ salts was used with a focus on the impact of the anions as refolding enhancers as well as different model proteins, among them recombinant plasminogen activator rPA. It was found that IL co-solvent systems with an intermediate capacity to solubilize proteins are effective as refolding additives, *i.e.*, were found to permit effective renaturation of rPA. The most promising refolding enhancer in this respect was [C₂mim][Cl]. By means of fluorescence correlation spectroscopy Sasmal *et al.* found that the presence of 6 M guanidine hydrochloride helps a protein, e.g., human serum albumin, to refold after addition of 1.5 M [pmim][Br] [46].

An interesting contribution quantifying the biomolecular interactions of the functional groups in proteins with biocompatible ILs, e.g., diethylammonium and triethylammonium acetate, has been reported recently [47]. Herein, transfer free energies (ΔG'_{tr}) of compounds such as Gly, Gly-Gly and cyclo(Gly-Gly) from water to aqueous ionic liquid solutions were determined depending on the IL concentration. The solubility in aqueous solutions decreased with increasing IL concentration (salting-out

effect). This was attributed to the non-favorable interactions between IL and individual amino acid residue(s). However, overall all ILs could stabilize the structure of the model compounds—with triethylammonium acetate being the strongest stabilizer among the ILs used. On the other hand, a negative contribution (ΔG^{tr}) of the peptide backbone from water to higher concentration of e.g., TMAA (50–70%) indicated that the interactions between the solvent and backbone unit are favorable.

The first attempts in peptide bond formation by enzymes, however, were already reported by Erbeldinger *et al.* in 2000 [31]. Here, the synthesis of Z-aspartame in a thermolysin-catalyzed reaction in $[\text{C}_4\text{mim}][\text{PF}_6]$ was demonstrated. Upon addition of 5% of water a biphasic reaction mixture was formed, and water exhibited a stabilizing effect on thermolysin. Also, the activity of the enzyme was dependent on the water content. In addition, an increased solubility was observed for the substrate and product upon addition of water. The reaction was monitored by HPLC after mixing of the reaction mixtures with 60% acetonitrile/water containing 10 mM orthophosphoric acid prior to injection. For the final procedure, however, a step allowing for the recycling of the IL was suggested after the reaction occurred. The yield of the product was 95%, *i.e.*, the same as found for organic solvents, e.g., ethyl acetate/water mixtures.

In 2007, Xing *et al.* reported the use of an α -chymotrypsin-catalyzed synthesis of protected tripeptides, e.g., the fragment Z-Tyr-Gly-Gly-OEt of Leu-enkephalin, in six different 1-alkyl-3-methylimidazolium hexafluorophosphates and tetrafluoroborates [30]. Again, it was found that the water content has a great influence on the two main factors, enzyme activity and coupling yield. At higher water content, the product is hydrolyzed, thus reducing the reaction yield. However, a small amount of water was essential for the reaction to succeed, *i.e.*, for hydrophobic $[\text{MOEMIM}][\text{PF}_6]$ the minimum content was 3%, while for the more hydrophilic $[\text{MOEMIM}][\text{BF}_4]$ 8% turned out to be favorable. The authors also reported that $[\text{MOEMIM}][\text{PF}_6]$ was the most suitable IL among the ILs tested, and peptides were prepared in 68–75% yield. Further enzyme-supported reactions in ILs have been reported by Wehofsky *et al.* [48] and are in detail in section 2.2.3, because larger peptide molecules prepared by ligation reactions are in the focus of that study. However, enzymatic acylation of smaller peptides has also been described by other groups, as can be exemplified by the lipase-catalyzed acylation of Lys-Ser x HCl with oleic acid carried out in either organic media (2-methyl-2-butanol) or an ionic liquid ($[\text{C}_4\text{mim}][\text{PF}_6]$) [49]. Substrate conversion in this case was found to depend on peptide solubility which was improved in the IL. Primarily chemoselective acylation occurred at the N_ϵ -amino group of lysine, while the side chain acylated serine derivative was not detected at all. Another example came from Malhotra *et al.*, who reported on the immobilized protease-catalyzed synthesis of peptides, e.g., Boc-Leu-Trp-OEt, in the ionic liquids $[\text{C}_4\text{mim}][\text{PF}_6]$, $[\text{C}_4\text{mim}][\text{BF}_4]$, and $[\text{C}_2\text{pyr}][\text{BF}_4]$ [50]. Here, product yields were comparable to those obtained in organic solvents, *i.e.*, a maximum yield of 49% for Boc-Leu-Trp-OEt was obtained.

In summary, 1,3-disubstituted imidazolium-based ionic liquids were successfully applied as additives to enzymatic syntheses. They were described to have a positive effect on enzyme-catalyzed coupling reactions, to increase the enzyme's activity, the coupling yields, the solubility of reagents as well as the product. Interestingly, the necessity of water for optimal coupling yields has been demonstrated regularly. Obviously a minimum amount of water (3–8%) is favorable for the activation respective enzyme.

2.1.2. Chemical Synthesis of Peptides

2.1.2.1. Condensation of Free Amino Acids in Solution

In a symposium report, Smith *et al.* demonstrated for the first time the peptide bond formation, e.g., for the generation of cyclopeptides, in ionic liquids at high temperatures (>100 °C). The thermally stable 1-butyl-3-methylimidazolium hexafluorophosphate was used, while water was absent in the reaction mixture. Surprisingly, at temperatures below 100 °C no reaction occurred [32]. Subsequent works also reported the condensation of amino acids in ILs, in particular in 1-butyl-3-methylimidazolium hexafluorophosphate [15,33,34,51]. Here, high coupling yields were achieved for quaternary α -amino acids that are much more difficult to couple than tertiary amino acids [15]. The reaction is thought to proceed through the stabilization of a charged species in the ionic solvent. Interestingly, tetra-, octa- and cyclooctapeptides were obtained in high yields (80–90%), which was comparable to organic solvents (e.g., THF). However, the crude peptides had much higher purities compared to classical organic synthetic protocols. A stabilization effect of the coupling reagents HATU and BOP in the ionic liquid was suggested to be the cause of this observation [34]. Still, finding a suitable extraction protocol for separating the product from the reaction mixture provides rather difficult, but in another publication good isolated yields (66–87%) were obtained by extracting of a protected dipeptide (Z-Gly-MPGOMe) [33,51]. In addition to these reports, Petiot *et al.* described the use of a new ammonium-based poly(ethyleneglycol)-based ionic liquid (PEG-IL) and its use for the synthesis of various dipeptides containing Phe, Gly, Ala, Leu and Glu under microwave activation [52]. As an example, the formation of Boc-Phe-Phe-OMe by using HATU as coupling reagent (3.3 equiv.) in a bis-PEG₃₅₀-IL carried out at 65 °C for 2 h yielded 64% of the product. It was discussed that the polar nature of bis-PEG₃₅₀-ILs may have a beneficial effect on the solubilization of the starting material, *i.e.*, amino acids and coupling agents. In addition, higher purities of the dipeptides obtained after work up, and reduced reaction times compared to previously reported peptide syntheses in ionic liquids were observed in this study.

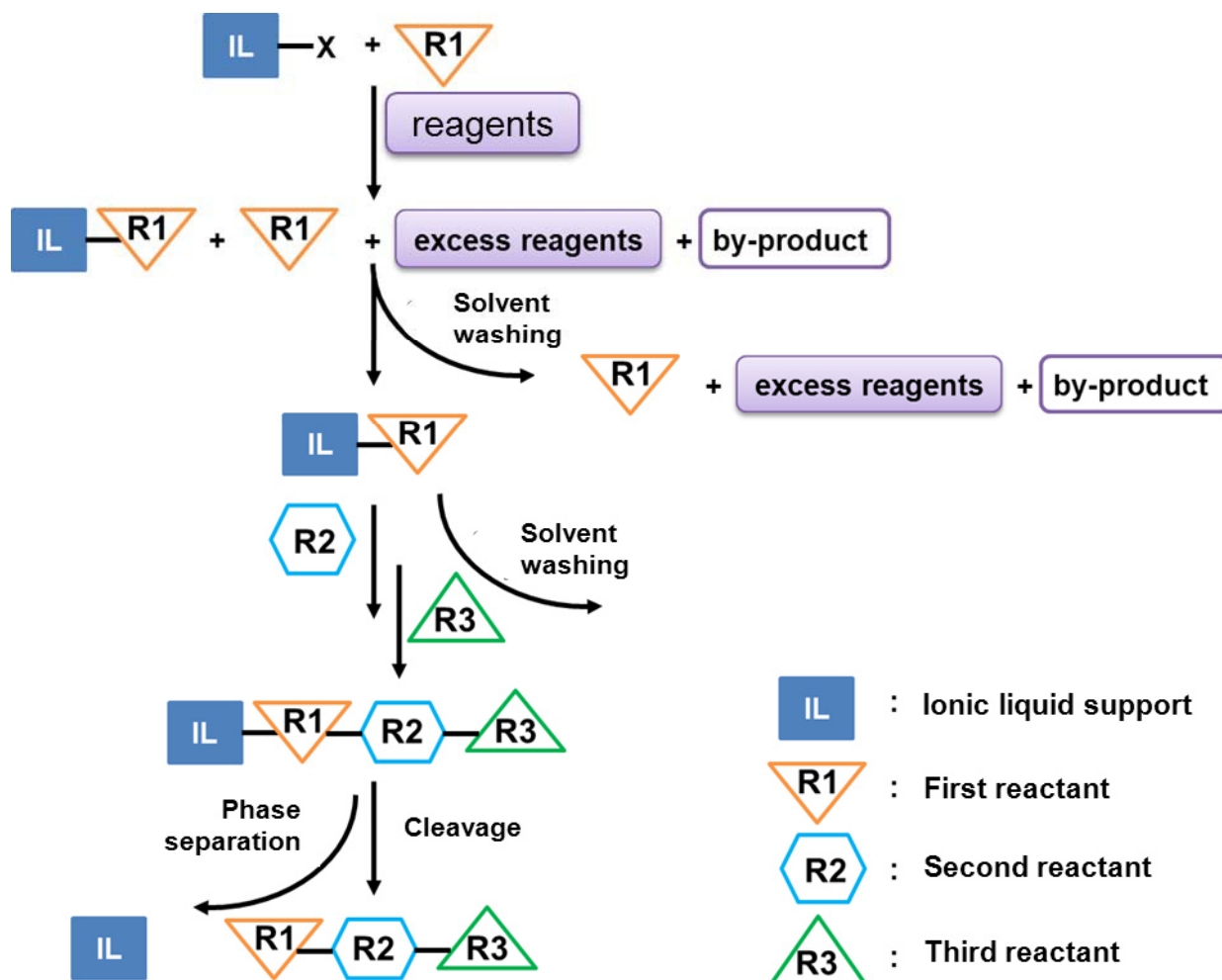
As demonstrated, first attempts have been undertaken to synthesize small peptides in ionic liquids as reaction media. In addition, the idea of a direct coupling of an amino acid to form a peptide bond by simply dissolving it in an ionic liquid and thus avoiding protection/deprotection steps seems promising for future applications of ILs in peptide synthesis [15].

2.1.2.2. Ionic Liquid-Supported Amino Acid Condensation

Solid phase peptide synthesis (SPPS) may have several limitations, e.g., solvation of reagents, slow coupling rates, difficult access of the reagents to the reaction sites, and nonlinear reaction kinetics. As a consequence, “liquid phase” methods were developed and applied recently (section 2.1). One of these methods represents the IL-supported peptide synthesis introduced by Miao and Chan in 2005 [36]. Here, substrates were directly attached to the respective IL, *i.e.*, 3-hydroxyethyl-1-methylimidazolium tetrafluoroborate, and by-products and excess reagents were removed from the desired product and intermediates by using simple washing steps (Figure 2). This synthetic procedure was used to produce Leu⁵-enkephalin of the sequence Tyr-Gly-Gly-Phe-Leu-OH. There were several advantages of this method, including (i) only a low excess of the coupling reagent (two equivalents PyBOP or DCC) was

required for the coupling steps instead of commonly used four to five equivalents, (ii) intermediates were easily purified by washing steps (liquid/liquid phase separation), and (iii) purity determination and structure identification were possible through spectroscopic methods (MS, NMR).

Figure 2. Scheme of ionic liquid-supported peptide synthesis (modified from [36]).



Furthermore, quantitative condensation of the first amino acid with DCC/DMAP and the second residue by using PyBOP/DIEA was achieved [36]. Additionally, IL-imidazolium attached oligomers were used as a soluble/solid support for the synthesis of hexapeptide building blocks. As an example/model the natural peptide Mucin4 containing a high number of serine and threonine residues was chosen. In contrast to previous studies, a mixture of HOBt/HBTU was used as coupling reagents, because of the difficulties to remove triperidinophosphine oxide from PyBOP during purification steps. Purification of the peptides was carried out by simple washing steps using water, because peptides were insoluble in water, while the soluble ionic supports could be washed out easily [53]. Shortly after this report, the same research group outlined the use of IL-attached oligomers for the large scale production of oligopeptides, oligosaccharides and oligonucleotides in a patent [35]. Other groups functionalized the ILs, namely cations 1-methylimidazolium and pyridinium and anions bromide, chloride, iodide and tetrafluoroborate, with amino, alcoholic or carboxylic groups. Rather high loadings (2.5–5.0 mmol/g) of the substrates of the ILs were obtained in these studies [54,55].

In our opinion, the coupling of amino acids in ILs can obviously be carried out at a high rate without the use of a large excess of coupling reagents, and this, in turn, is one big advantage of the use of ILs in peptide chemistry.

2.2. Peptide Modifications

In this part we describe the progress made with respect to reactions other than the general peptide assembly methods, e.g., the synthesis of radiolabeled peptides, the formation of disulfide bridged polypeptides by oxidative folding and different ligation procedures used to obtain oligopeptides and small proteins. This novel and, at the current state, sparsely investigated research field hints to the successful use of ILs for such processes and opens exciting new ways for the optimization of protocols that are hampered by many drawbacks of conventional methods applied thus far.

2.2.1. Radiolabeling of Peptides

Radiolabeling is an established technique with many applications in biological research, such as the investigation of the distribution of labeled markers *in vivo* (e.g., in humans and in animals) by positron emission tomography [60]. The synthesis of the radiolabeled bioactive compounds, in particular peptides, is usually time consuming and requires complicated procedures and special equipment to purify the product from unwanted non-radioactive and radioactive byproducts. Strictly anhydrous conditions are highly recommended, for example, for fluoride ion displacement reactions in the conventional protocols.

A new approach in this respect has been proposed by Kim *et al.* in 2003 (Figure 4) [37]. Herein, fluorine-18 labeling of alkyl mesylate or alkyl halides, e.g., 2-(3-methanesulfonyloxypropoxy)-naphthalene, in ILs containing a hydrophobic cation ($[C_4mim][X]$, where $X = [BF_4]^-$, $[PF_6]^-$, $[SbF_6]^-$, $[OTf]^-$, $[NTf_2]^-$) was described for the first time. This method was successful despite the addition of a small amount of water, which is in contrast to the conventional method used. The ILs enhanced the reactivity of fluorides and also, a reduction of by-product formation (alkenes, alcohols) was observed. This approach led to further developments in this field as reported by Schirmacher *et al.* [38]. The IL $[C_4mim][OTf]$ was useful for the labeling of short model peptides, e.g., GSH [38]. Thus, radiolabeling performed in IL (1,3-dialkylimidazolium salts) obviously occurs without the need for complicated procedures and strictly anhydrous conditions, but by direct addition of the IL to the reaction mixture and further simple purification steps.

Figure 4. Radio-fluorine-18-labeling performed in an ionic liquid (modified from [37]).



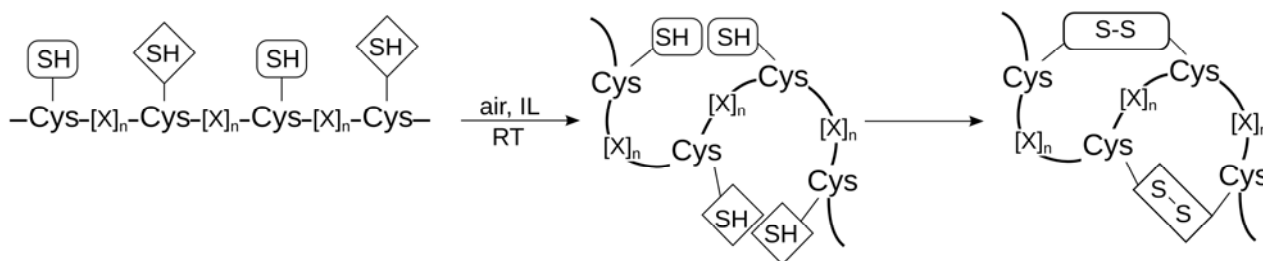
2.2.2. Disulfide Bond Formation and Peptide Folding

Establishing efficient strategies for the synthesis of cysteine-rich peptides has always been a challenging task in the past. Generally, the incorporation of cysteine residues in peptide sequences is difficult as it requires the protection of the highly reactive thiol group [18,61–63]. Cysteine, however,

is very important since many naturally occurring peptides contain intramolecular disulfide bonds that stabilize the biologically active conformation [18,64,65]. Thus, one application of cysteine-residues in peptide chemistry represents disulfide bridge formation that may be achieved at various stages of the synthesis, on-resin as well as in solution [62,66–68]. Protecting group strategies differ according to the number of cysteine residues and disulfide bonds to be formed. While peptides containing one disulfide bridge are mainly prepared by oxidation (e.g., air, H_2O_2 , iodine) in solution at low concentrations (10^{-3} to 10^{-4} M), for peptides with two or more disulfide bridges the situation is more complicated since intramolecular mispairing (scrambling) and/or intermolecular dimerization or oligomerization may occur [61–63].

The major disadvantages of these conventional methods are: (i) the requirement for a high dilution of the reaction mixture, (ii) very long reaction times, (iii) inadequate solubility of hydrophobic peptides in buffer solutions (including the need for organic solvents: isopropanol, methanol, acetonitrile), and (iv) the difficulty to control oxidation that requires the reaction to be carried out at low temperatures (usually 4 °C). Consequently, there are huge problems for producing such peptides in large amounts and sufficient purity. Over the years, many scientists were searching for a method that can overcome these problems. In our own work [14], we recently described for the first time the use of 1-ethyl-3-methylimidazolium-based ILs with different anions including small hydrogen-bond acceptors (acetate [OAc^-], diethyl phosphate [DEP^-]) as well as larger, less-coordinating anions (tosylate [OTs^-], dicyanamide [$\text{N}(\text{CN})_2^-$]) for the oxidative folding of neuropeptides (Figure 5).

Figure 5. Oxidative folding of cysteine-rich peptides performed in ILs.



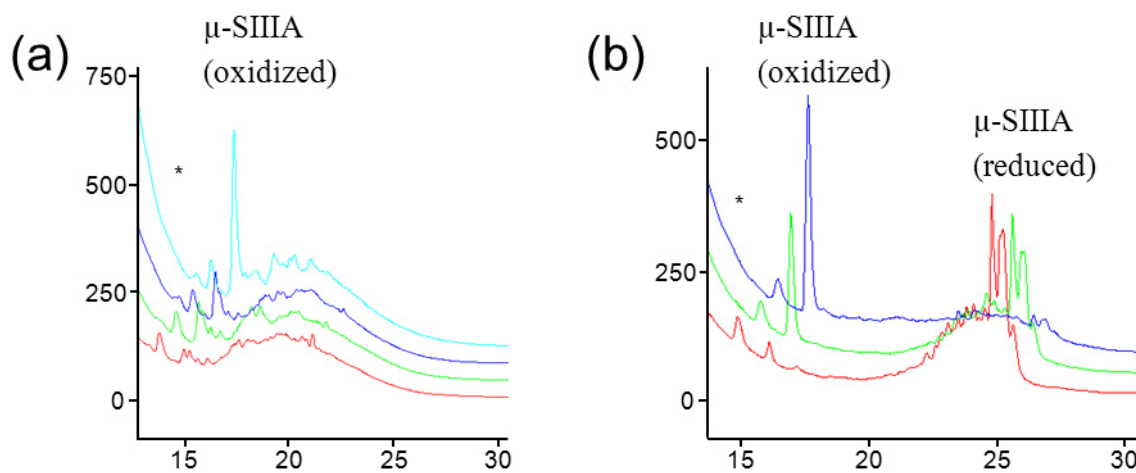
A very good solubility of hydrophobic peptides, e.g., δ -EVIA and δ -SVIE, containing six cysteine residues was observed in these ILs. The reaction was performed in ILs of low water content (between 0.5–3%), in a very small volume and a rather high concentration of the peptides (10–15 mM) [14]. Thus, this seems to be a promising method for up-scaling the production of cysteine-rich peptides, also because the formation of by-products and misfolded species was minimized in comparison to conventional methods (oxidation in buffer solution [69,70]). In contrast to the latter, there is no need for additional redox reagents, thus the reaction can be performed in the pure IL. A stabilization of pre-formed secondary structures of the peptides investigated by the ILs was suggested to be the cause of these observations.

Later, different naturally occurring cysteine-rich peptides were studied to understand the nature of interaction between ILs and peptides. Various model peptides—short (10 amino acids) and medium-sized (20–40 amino acids) native peptides with varying numbers of cysteine residues in their structure (from two to six) were subjected to these investigations. Depending on the net charge, amino

acid content and hydrophobicity of the peptide, the efficiency in the ILs used differed strongly. In contrast to conotoxins μ -SIIIA, μ -PIIIA, δ -EVIA, and δ -SVIE that showed best results when using $[\text{C}_2\text{mim}][\text{OAc}]$ as reaction medium [14], optimum conversion of linear CCAP-vil [39] was observed in $[\text{C}_2\text{mim}][\text{OTs}]$ (yield >90% according to HPLC analysis) and to a lesser extent in $[\text{C}_2\text{mim}][\text{DEP}]$ (yield >80% according to HPLC analysis). Interestingly, the target peptide did not form the oxidized version in $[\text{C}_2\text{mim}][\text{OAc}]$ in acceptable yields (57%), and the yield was even lower in case of $[\text{C}_2\text{mim}][\text{N}(\text{CN})_2]$ (approximately 44%). When increasing the lipophilicity of the ionic liquids (by increasing the alkyl substituent from $[\text{C}_2\text{mim}][\text{OAc}]$ to $[\text{C}_4\text{mim}][\text{OAc}]$), the yield of the oxidized product decreased to approximately 47% [39].

It is clear from these studies that oxidative folding depends on the proper choice of the ionic liquid. Also, peptides with a high net charge are apparently more efficiently oxidized in ILs containing more basic anions, such as $[\text{C}_2\text{mim}][\text{OAc}]$, compared to small uncharged peptides such as CCAP-vil, where the most suitable ionic liquid contains a less basic anion, *i.e.*, $[\text{C}_2\text{mim}][\text{OTs}]$. Unpublished data related to our former work showed interesting dependencies of the reaction outcome and product purities on water content and reaction temperature (Figure 6). This can be exemplified for the conopeptide μ -SIIIA, where already an increase of the water content from 3–10% was disadvantageous for product formation. In contrast, an increase in temperature up to 80 °C was favorable for the oxidation reaction.

Figure 6. RP-HPLC elution profiles of μ -SIIIA (crude mixtures) after oxidative folding in $[\text{C}_2\text{mim}][\text{OAc}]$ depending on (a) water content (cyan with 3% water, blue—5%, green—7.5%, red—10%) and (b) temperature (blue curve—80 °C, green—60 °C, red—40 °C).
* Peak from ionic liquid.



The role of ILs in oxidative folding of peptides is still sparsely investigated, since further candidates are necessary to derive a scientific explanation for these observations, yet the application of ILs for this reaction clearly has the potential to be developed as a key technology in peptide chemistry in the future. However, an interesting aspect with respect to the folding behavior of peptides was recently presented by Huang *et al.*, who reported on the investigation of the folding of model peptides in neat ionic liquids by CD spectroscopy [71]. This study was motivated by the fact that previous IL studies have almost exclusively focused on large functional proteins, such as lysozyme, ribonuclease A and human serum albumin. Indeed, the authors were correct in stating that the interpretation of the results

obtained for proteins are complicated by their size and the different degrees of freedom available to either a peptide or a protein. Thus, they focused on small peptides, e.g., α -helical AKA₂, miniprotein Trp-cage (contains e.g., α -helical part and 3₁₀ helix), and Trpzip4 (β -hairpin), in order to understand the effect of ILs, in this case [C₄C₁pyrr][NTf₂], on well-designed secondary structures. The model sequences used represent optimized native folds in an aqueous environment. Huang *et al.* characterized the thermal transitions of these peptides using circular dichroism (CD) spectroscopy and found that, in contrast to the aqueous data, the far-UV CD spectra of AKA₂ and Trp-cage in [C₄C₁pyrr][NTf₂] indicated only minimal helical structure at low temperatures. However, in both cases, this data and the temperature-dependent far-UV CD spectra suggested apparent heat-induced structure formation with structure persisting to the highest temperatures recorded (92 °C). In contrast to the results of the helical peptides, the CD data for Trpzip4 indicated huge destabilization by the IL medium with increasing temperature [71]. In addition to this report, Debeljuh *et al.* studied the impact of triethylammoniummesylate (TeaMS) on the structure of the Abeta(1–40) peptide for Alzheimer's disease [72]. They found that the IL was able to induce a conformational change and also that this structural change influences the self-assembly of the peptide into amyloid fibrils. In our opinion, these studies are an important starting point for further investigations on peptide conformations and peptide folding observed in ionic liquids.

2.2.3. Ligation Reactions in Peptide Synthesis

Chemical ligation reactions are useful tools for the assembly of small to medium-sized proteins and protein domains starting from synthetic peptide fragments [73–75]. This type of chemical protein synthesis allows the modification of the molecular structure offering new insights into protein functions. Though 50 to 100-mer peptides can be assembled on a solid support, synthesis of longer peptides and proteins by the Merrifield method is often limited (see sections 2.1 and 2.1.2). Methods of segments ligation have been developed primarily in order to circumvent limitations in chain length. These segment condensation reactions, on the other hand, may be hampered by the low solubility of peptide fragments for the condensation reaction carried out in water or aqueous/organic solvent mixtures, and the stability of biomolecules. Both aspects are very important for the ligation reaction itself, and have an impact on later purification steps. One of the most widely used methods in this respect is the so-called “Native Chemical Ligation” (NCL) introduced by Kent and coworkers [73–75]. Here, an unprotected peptide with a thioester at the C-terminal and an unprotected peptide with an N-terminal cysteine react to yield a peptide bond. The mechanism of this reaction is shown in Figure 7. Related methods are the Staudinger ligation, which occurs between a peptide azide and a phosphinothioester, and protease-catalyzed ligations, where specific peptide esters (e.g., substrate mimetics) react regioselectively and stereospecifically with nucleophilic peptide species [73,76,77].

In section 2.1.1, we described examples of thermolysin- and α -chymotrypsin-catalyzed couplings of amino acids to generate short peptides. Bordusa and co-workers, however, highlighted the protease-catalyzed ligation using ILs for the first time in 2008 [48]. The reaction was performed in a mixture of buffer and 1,3-dimethylimidazolium dimethylphosphate [C₁C₁IM][Me₂PO₄] (40:60) (Figure 8). Several positive effects were observed for this reaction in the IL mixture in comparison to conventional methods, e.g., (i) the complete suppression of proteolytic side reactions (proteolytic hydrolysis and decreases competing hydrolysis activity of proteases toward the acyl donor esters to

some extent), (ii) high turnover rates, (iii) protease stability, and (iv) good solubility of the reactants and also of the reaction products.

Figure 7. Native chemical ligation occurs between two unprotected peptide segments: (i) transthioesterification; (ii) S to N acyl transfer [73,74].

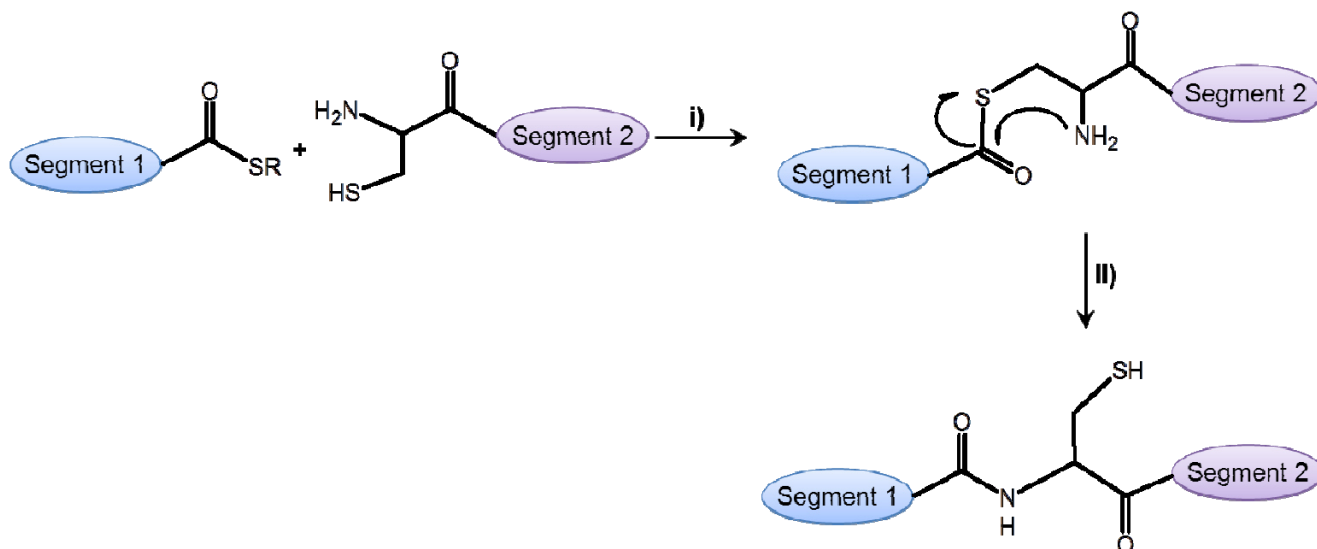
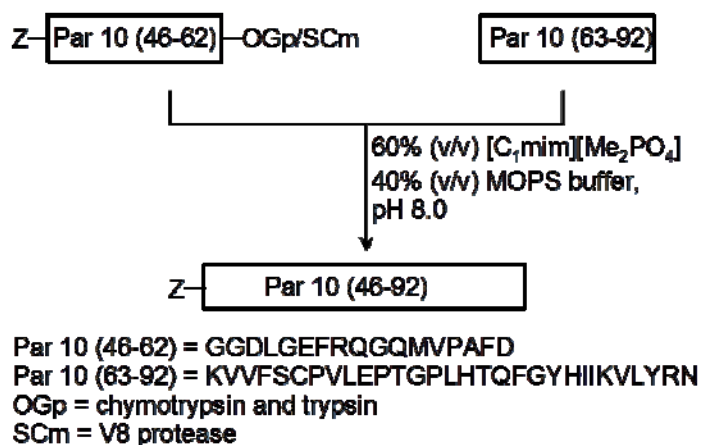
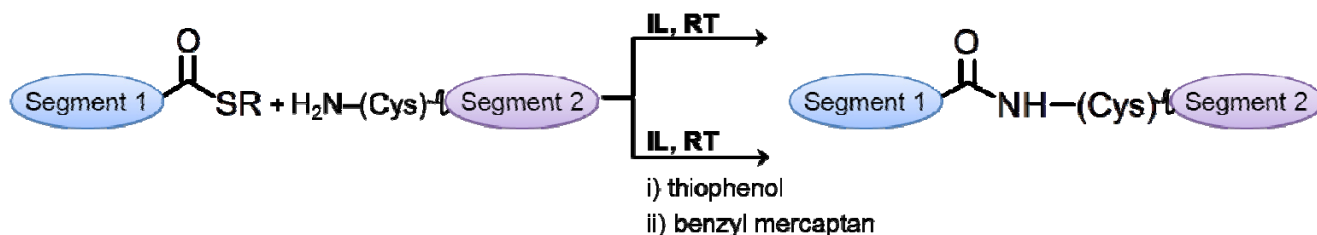


Figure 8. Protease-catalyzed ligation of parvulin (residues 46–92) employing peptidyl prolyl cis/trans isomerase (modified from [48]).



Recently, we reported on the application of native chemical ligation to produce a 66-mer anticoagulant cysteine-rich peptide (Tridegin) that has been shown to act as an inhibitor of factor XIIIa (Figure 9) [78]. Generally, NCL may be carried out with and without additives, *i.e.*, thiols (e.g., thiophenol, benzyl mercaptan) as catalysts [79–82]. Therefore, the comparison of segment ligation in buffer and ILs in the presence or absence of additives was performed as well. The results showed that NCL indeed yields the ligated product. The linear Tridegin was identified in [C₂mim][OAc] even in the absence of thiophenol and benzyl mercaptan. In addition, by using additives the ligated product was also formed in [C₂mim][DEP] that correlates well with the results published by Bordusa and co-workers mentioned above [48].

Figure 9. Native chemical ligation performed in ILs (modified from [78]).

In summary, the results of our and other groups demonstrate that medium-sized and large disulfide-bridged peptides are correctly formed in biocompatible ionic liquids. Moreover, the selection of peptide candidates examined thus far suggests that the correct intramolecular hydrogen bond interactions are not significantly perturbed by the presence of ILs, instead ILs seem to have a favorable effect on stabilizing pre-formed secondary structures of peptides.

3. Ionic Liquids for Peptide Purification and Characterization

Peptide purification and chemical characterization are the most important steps after a peptide has been synthesized. Usually peptides are purified using reversed-phase high performance liquid chromatography (RP HPLC) and characterized by different methods such as mass spectrometry (ESI or MALDI MS), thin-layer chromatography (TLC), and capillary electrophoresis. The structure of peptides can be determined by NMR spectroscopy. In this section, we give an overview of the applications of ILs in those methods, in which ILs have been applied so far.

3.1. Peptide Separation by Liquid Chromatography Methods

The use of ILs for chromatography reveals promising perspectives for the establishment of new separation methods and protocols. Examples for the application of ILs in separation techniques are well summarized in the review of Han *et al.* [83], but only few reports are found for the application in peptide chemistry. This field of investigation is rather new, however, the academic interest has increased enormously. Here, we explicitly highlight new applications of ILs for peptide separation and characterization, which were established in the last few years.

3.1.1. Reversed-Phase HPLC

HPLC is a widely used method for separation and purification of biomolecules. Especially for peptide analysis RP HPLC is the most common method applied. Some reports have been published about the application of ILs in RP HPLC for separation of nucleotides, geometric isomers, ephedrine or carboxylic acids [84–87]. These reports are based on the use of ILs as additives for the mobile phase [84] or as a stationary phase [87]. The commonly used ILs for these processes contained 1,3-dialkylimidazolium or *N*-alkylpyridinium cations and chloride, tetrafluoroborate, hexafluorophosphate, or trifluoroacetate as the anion depending on the analyte [86]. Nevertheless, none of the reports represents an example of an IL-supported optimal separation of peptides. Marszall *et al.* suggested that it is not possible to find a suitable IL for separation and purification processes [88] according to the principle, which can be

expressed by the basic chemical aphorism “*similis similibus solvuntur*” or “similar dissolves in similar”. Therefore, when the sample, *i.e.*, the peptide, is polar, the use of a polar IL with chaotropic anions, such as hexafluorophosphate or perchlorate was proposed, while a less polar IL with a cosmotropic anion (e.g., chloride) was recommended for a nonpolar analyte [86]. That is exactly why an individual protocol for different analytes is required.

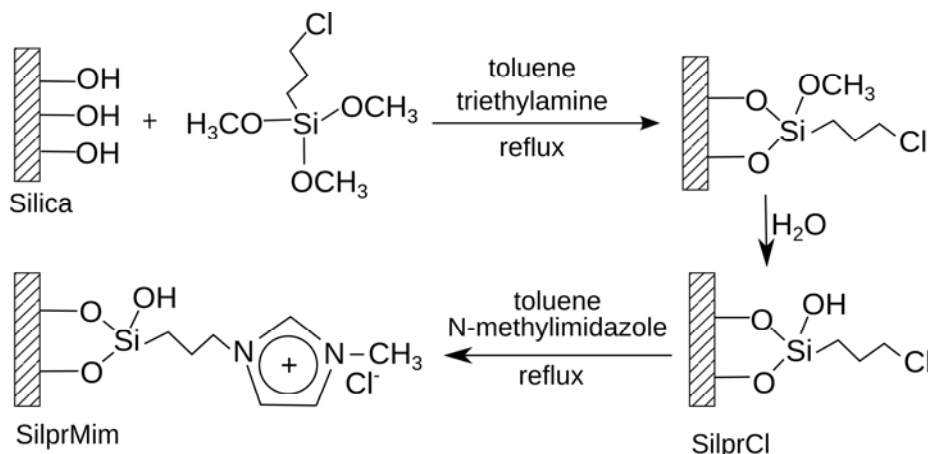
3.1.1.1. ILs as Additives for RP HPLC Mobile Phase

Ionic liquids have been used as additives to mobile phases in RP HPLC [86], because the relatively high viscosity of some ILs prevents their usage as pure eluent, in contrast to standard organic solvents (MeOH, ACN) commonly used for peptide purification. However, in the work of Polyakova *et al.*, IL concentrations of 2–50 mM resulted in an improvement in the peak shape and decreased peak tailing or band broadening [84]. They use the separation of *o*-, *m*- and *p*-amino benzoic acid involving [C₂mim][BF₄], [C₄mim][BF₄], [C₂mim][MeOSO₃] and [C₈mim][MeOSO₃] as additives. The best separation was found with the methylsulfate systems and a concentration of 1.0–8.0 mM/L. The kinetics of the solute-stationary phase is responsible for this feature. Later, this group also reported on changes in the retention time of analyte elution, which decreased or increased depending on the type of IL used, *i.e.*, on the solute-stationary phase interactions [85]. When the IL ions coat the silanol groups of the silica stationary phase the interactions of the analytes with the hydroxyl groups of the silica phase are minimized. Thus, the analyte eluted earlier [86]. Different models were proposed to explain this behavior: (1) ion-pairing, (2) adsorption of an IL layer on the C18 surface, or (3) linear solvation free energy relationship (LSER) [89,90]. The latter model was developed to describe the retention behavior in RP chromatography and included the calculation of intermolecular interactions [89,90].

3.1.1.2. ILs as Material for RP HPLC Stationary Phase

Similar to the use of ILs as mobile phase for RP HPLC, researchers tried to use ILs as stationary phase. For example, the ionic liquid cation was covalently attached to the silica substrates by means of a linker, e.g., 3-mercaptopropylsilane or propyl [86,91]. Such stationary phases were named “surface-confined ionic liquids (SCIL)” (Figure 10) [91,92].

Figure 10. Procedure for the preparation of a *N*-methylimidazolium-functionalized silica gel stationary phase (modified from [92]).

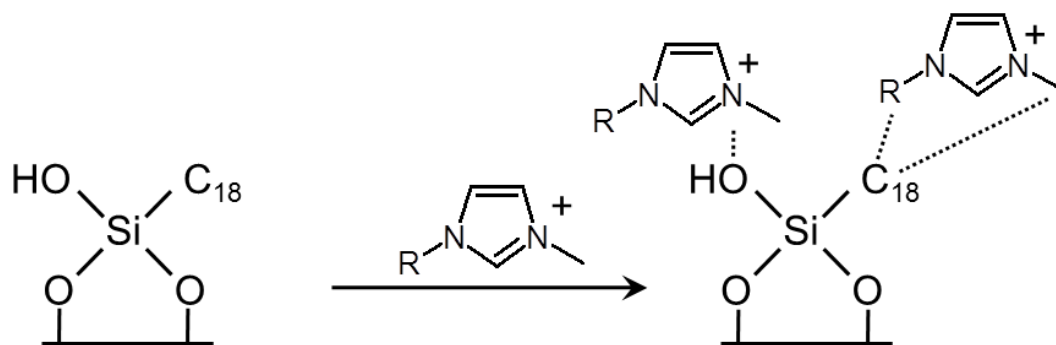


The commonly used ILs for this purpose are based on 1,3-dialkylimidazolium or *N*-alkylpyridinium cations with variably associated anions [91,93]. In 2010, the first example of SCIL applied to the separation of small peptides was reported by Chitta *et al.* [87]. They used a dipeptide (GY), a tripeptide (VYV), two pentapeptides (neuropeptides Leu-enkephalin, YGGPL, and Met-enkephalin, YGGPM), and an octapeptide (angiotensin II, DRVYIHPF) as subjects of investigation. The silica surface was modified with *n*-butylimidazolium bromide, and the effect of trifluoroacetic acid as ion-pairing agent was evaluated. In RP-HPLC of peptides, ion-pairing agents are commonly used as mobile phase additives. Such agents are primarily selected depending on the charge of the peptide(s) to be separated. Several volatile perfluorinated acids, such as TFA and pentafluoropropionic acid, are generally applied in this respect. They affect the retention behavior of a peptide through interactions of the ions with charged functional groups of the peptide's side chains, e.g., of basic residues Lys, Arg and His, as well as of the N-terminal amino group. In turn, hydrophilicity of the peptides is reduced, since net charge is reduced and thus, ion-pairing ions influence the hydrophobic character of a peptide resulting in increased hydrophobic interactions with the reversed-phase sorbent [87]. However, TFA in different concentrations (0.001–0.1% v/v) did not act as ion-pairing agent for the elution systems used by Chitta *et al.* and thus, can be substituted by formic acid in these systems. Although the SCIL phase possesses reversed-phase character, it was found that electrostatic interactions dominated at high organic and/or low pH modifier concentrations [87].

3.1.2. Thin Layer Chromatography

Ionic liquids were not only applied to RP HPLC, but also used for TLC and normal phase HPLC [94,95]. However, characterization of basic or amphoteric samples is challenging if using these methods. The problem is that acidic silanol groups of the stationary phase interact strongly with organic bases (Figure 11) resulting in (i) an increase in retention, and (ii) a broadening and/or tailing of sample peaks [96].

Figure 11. Schematic representation of a silica-based stationary phase interacting with an imidazolium-based IL, where -R represents alkyl substitutions (proposed by He *et al.*) (modified from [96]).



It has already been reported that the use of IL additives may help to overcome these problems. Ionic liquids block silanol groups and in this way provide the desired separation of basic drugs as test analyte [97]. Similarly to RP HPLC, alkyl substituted imidazolium ILs were used for TLC and HPLC

as well. Baczek *et al.* investigated the effect of the percentage amount (0–10%) of the ILs as additives to the eluent on the separation of peptides by TLC. For approximately 50 peptides the R_f values increased with an increased concentration of 1,3-dimethylimidazolium methyl sulfate in a water/0%–5% acetonitrile/0.1% TFA mixture [98]. The influence on the retention of the peptides with an increasing IL concentration was weaker with increasing length of the amino acid sequences. In the study of Baczek *et al.*, 1-ethyl-3-methylimidazolium tetrafluoroborate was applied. Upon addition of acetonitrile (1.5%) a retention coefficient with third-degree polynomial function was derived, however, a quadratic function was needed in case of the experiment without acetonitrile. This was explained by the interaction of the IL with the stationary phase [99].

Another method involved the addition of IL and matrix (α -cyano-4-hydroxycinnamic acid) to the separation mixture. This optimized the chromatographic method allowing for a direct coupling of the thin-layer chromatograms with imaging by MALDI mass spectrometry (see also section 3.2) [100].

3.2. Mass Spectrometry

MALDI mass spectrometry is a widely used method for the characterization of biomolecules [101,102] as individual components from mixtures or even whole tissues [16]. The main characteristic of this method is the “soft” ionization in comparison to other methods, e.g., electrospray ionization [103].

For successful analytical results a suitable matrix is indispensable. Matrices absorb the energy of the laser and support ionization of the analyte. They are usually small acidic compounds that co-crystallize with analyte molecules and provide the access of protons [15,104]. Main problems of standard solid matrices are low homogeneity, formation of so called “hot spots” and/or adducts of the analyte with cations and anions. This results in low measurement reproducibility, decreased sensitivity, broadening or suppression of peaks [105]. Commonly used matrices for MALDI MS of peptides and proteins are α -cyano-4-hydroxycinnamic acid (CCA), sinapinic acid (SA), 2,5-dihydroxybenzoic acid (DHB), and indole acrylic acid.

ILs as matrices for peptide analysis with MALDI mass spectrometry were first reported in 2004 by Mank *et al.* [106]. Different combinations of matrices were probed, but the CCA-butylamine (1:1) was described to be the best choice for MALDI MS of the peptides investigated. One of the reasons was that alkali adducts with the peptides were observed if the DHB-butylamine (1:1) matrix was applied, which was prepared in an ethanolic ionic liquid 1:1 (v/v) mixture. Herein, it was shown for the first time that this DHB-IL matrix can be used for a direct and easy screening of enzymatic reactions [106]. In the same year, another report of Zabet-Moghaddam *et al.* was published, where equimolar mixtures of SA, CCA, and DHB with organic bases, such as tributylamine, 1-methylimidazole or pyridine were used [105]. It was shown that ILs were not suitable as MALDI matrices without addition of classical matrices. The best results were achieved for the use of water immiscible ILs, e.g., [C₄mim][NTf₂] or [C₄mim][PF₆], with the addition of CCA. At the same time, peptides behaved differently in the IL-matrix compared to the IL-free medium [105]. This fact was also mentioned in later research works [103], stating that the homogeneity of the samples decreased in comparison to conventional matrices [105]. However, these first examples on the use of IL-mixtures with classical matrices for peptide analysis demonstrated an improved spot-to-spot reproducibility and better ionization efficiency. Since then, numerous reports appeared in the literature aiming to find out the best matrix

system for peptide analysis. Many different combinations of the basic part of an IL and the acidic part of the standard matrices were used [107]. From that time on, these mixtures are called “ionic liquid matrices (ILM)” [103]. Jones *et al.* also showed an enhanced sensitivity of the MALDI TOF method for peptide characterization if such ILM were used [103]. However, the first reports of the analysis of synthetic peptides with ILMs were described in 2006. A reduction of peptide adduct and matrix cluster formation, an improvement of the signal-to-noise ratio and an increase of the peak intensities was observed for the use of a CCA-pyridine ILM (2:1, v/v). Also, the detection limits were reduced using the described ratio [108]. In addition, an improvement in sample homogeneity was reported by Tholey *et al.* [109] who showed that the quantitative analysis of peptides was possible without the use of an internal standard. Additionally, a linear correlation between the peptide amount and the signal intensities was observed. Here, the samples were prepared with the ILM in a matrix(acid)-IL(base) ratio of 1:1. As bases, *N,N*-dimethylethylenediamine (DMED) and 3-(dimethylamino)-1-propylamine were used, and as matrix CCA and indole acrylic acid were applied. Both matrices were found to be suitable for peptide analysis. The addition of a matrix additive (1% H₃PO₄) to increase the detection limit and signal intensities was reported by the same research group [110]. Phosphopeptides were used as targets for the MALDI MS analysis described.

A new way of sample preparation for high sensitivity quantitative analysis of peptides was proposed by Palmblad *et al.* [111]. A conductive hydrophobic TeflonTM tape was used as sample surface in the analysis in order to concentrate the analyte to microliter range and to achieve better ionization. Also, the successful attachment of cations (Cu²⁺ and K⁺) to target peptide angiotensin I in the ILM CCA/3-aminoquinoline/glycerol (1:4:6) in comparison to the standard dried droplet method and powder mixture sample preparation was reported by Hortal *et al.* shortly after [112]. In 2008, a further ILM was proposed as suitable matrix for peptide analysis with MALDI mass spectrometry, which was a mixture of CCA and aniline. This matrix was also structurally characterized. A similar or even improved signal-to-noise ratio was observed in comparison to the pure CCA matrix [113]. Other ILMs such as *N,N*-diisopropylethylammonium-CCA and *N*-isopropyl-*N*-methyl-1-butylammonium-CCA matrices were reported to be very suitable for the analysis of peptides and proteins in the molecular weight range from 1 kDa to 270 kDa. For example, these matrices showed much higher intensities of the analyte [114].

The efforts to use ILMs for the analysis of glycopeptides and glycans succeeded when using the matrix [1,1,3,3,-tetramethylguanidinium-2,4,6,-trihydroxyacetophenone] (GTHAP). The detection of glycopeptides containing large carbohydrate groups and small peptide chains was higher than in the solid matrix (2,4,6,-trihydroxyacetophenone). In addition, the degree of glycopeptide decomposition was suppressed if using GTHAP [115]. The avoidance of sugar fragmentation as well as the improvement of homogeneity and reproducibility of the spots were also important issues in the report of Gimenez *et al.*, who investigated the effectiveness of different ILMs with SA and DHB on the ionization of intact glycoproteins with several degrees of glycosylation [116]. All ILMs preparation procedures used butylamine as organic base and were prepared either with or without an acid, e.g., 0.1% of TFA or 1% of H₃PO₄. The authors demonstrated that the chemical composition of an ILM strongly influenced the analysis of their test substances. In addition, ionization efficiencies and spot homogeneity were best if using ILMs with higher amounts of the organic salt, *i.e.*, the ILM containing 3 M of butylamine and an equimolar amount of either DHB or SA was the optimal preparation

procedure. The results for the average molecular mass values also indicated that sugar fragmentation was prevented for either mixture, while the SA-ILM was best with respect to reliable analysis, sensitivity, sharpness of ion peaks and application of a wider range of laser intensities [116]. Finally, new reports by Fitzgerald and co-workers showed that ILMs are suitable for MS imaging, e.g., of onion skin, as well [117]. Here, 1-methylimidazolium-CCA and tripropylammonium-CCA were applied as matrices. Both were suggested to improve the detection limits to about 30- to 40-times in comparison to the standard matrix. Again, homogeneity and stability of the sample surface during the crystallization process was improved [117].

One further application of ILs in mass spectrometry other than the MALDI technique should be mentioned here as well, where ionic liquids were used for the identification of integral membrane proteins (IMPs) by microcolumn reversed phase liquid chromatography (μ RPLC)-electrospray ionization tandem mass spectrometry (ESI-MS/MS) [118]. The authors here used 1% (v/v) [C₄mim][BF₄] in NH₄(HCO₃) buffer (pH 8.3) as the sample preparation buffer for IMPs analysis. Furthermore, compared to the commonly used methods (e.g., sodium dodecyl sulfate and methanol) a three times higher number of identified IMPs from rat brain was obtained, which was explained by the fact that [C₄mim][BF₄] may provide higher solubilizing ability for IMPs and good compatibility for tryptic digestion.

3. Conclusions

The application of ionic liquids in peptide chemistry is currently in the exciting phase of its development. The assembly of peptides by SPPS or in solution using ILs has been described, but primarily for shorter representatives. Therefore, these investigations are, in our opinion, in an early stage of development. Nevertheless researchers continue to find optimal more simple and straightforward conditions to carry out reactions involving amino acid derivatives and peptides. In general, synthesis and characterization of peptides using ionic liquids seems to have several positive effects, including the better solubility of solutes, the use of higher temperature, the possibility to reduce the concentrations of excessive agents, higher reaction yields, and higher resolution (analytics).

It has been shown for synthetic processes that ILs are not just simple solvents, but may act as a functional part in a reaction process, e.g., by directly interacting with analytes. In addition, many reports concerning the characterization of peptides by mass spectrometry, in particular MALDI-TOF MS, and different types of chromatography were published in the last decade. On the other hand, methods such as peptide/protein structure determination by NMR spectroscopy or separation by capillary electrophoresis are not as popular in the literature, yet are starting to gradually emerge. The reason for this may be that a deeper understanding of the interactions of IL cations and anions with specific side chains and/or the peptidic backbone in dependence on the IL environment and the absence and/or presence of water is still missing. Thus, a closer examination of these interactions would support both, the design of suitable ILs for special applications concerning the generation of a peptide (primary amino acid sequence) and/or its conformation (three-dimensional structure) and the optimization of purification and characterization protocols. Future studies will therefore focus on intensive and systematic investigations in order to address the questions raised by the early examples summarized in the present report.

Acknowledgments

We thank Miriam Böhm for support with the manuscript. This work was financially supported by the Deutsche Forschungsgemeinschaft within priority program 1191 (to A.S. and D.I.) and the Friedrich Schiller University of Jena (to D.I.).

Reference

1. Sheldon, R. Catalytic reactions in ionic liquids. *Chem. Commun.* **2001**, 2399–2407.
2. Welton, T. Room-temperature ionic liquids. Solvents for synthesis and catalysis. *Chem. Rev.* **1999**, *99*, 2071–2083.
3. Wasserscheid, P.; Keim, W. Ionic liquids—New “solutions” for transition metal catalysis. *Angew. Chem. Int. Ed.* **2000**, *39*, 3772–3789.
4. Lange, C.; Patil, G.; Rudolph, R. Ionic liquids as refolding additives: *N'*-alkyl and *N'*-(omega-hydroxyalkyl) *N*-methylimidazolium chlorides. *Protein Sci.* **2005**, *14*, 2693–2701.
5. Moon, Y.H.; Lee, S.M.; Ha, S.H.; Koo, Y.M. Enzyme-catalyzed reactions in ionic liquids. *Korean J. Chem. Eng.* **2006**, *23*, 247–263.
6. Hekmat, D.; Hebel, D.; Joswig, S.; Schmidt, M.; Weuster-Botz, D. Advanced protein crystallization using water-soluble ionic liquids as crystallization additives. *Biotechnol. Lett.* **2007**, *29*, 1703–1711.
7. Baker, S.N.; Brauns, E.B.; McCleskey, T.M.; Burrell, A.K.; Baker, G.A. Fluorescence quenching immunoassay performed in an ionic liquid. *Chem. Commun.* **2006**, 2851–2853.
8. Kragl, U.; Eckstein, M.; Kaftzik, N. Enzyme catalysis in ionic liquids. *Curr. Opin. Biotechnol.* **2002**, *13*, 565–571.
9. Stark, A. Ionic liquid structure-induced effects on organic reactions. *Top. Curr. Chem.* **2009**, *290*, 41–81.
10. Fujita, K.; MacFarlane, D.R.; Forsyth, M. Protein solubilising and stabilising ionic liquids. *Chem. Commun.* **2005**, 4804–4806.
11. Baker, S.N.; McCleskey, T.M.; Pandey, S.; Baker, G.A. Fluorescence studies of protein thermostability in ionic liquids. *Chem. Commun.* **2004**, 940–941.
12. Feher, E.; Major, B.; Belafi-Bako, K.; Gubicza, L. On the background of enhanced stability and reusability of enzymes in ionic liquids. *Biochem. Soc. Trans.* **2007**, *35*, 1624–1627.
13. Dreyer, S.; Kragl, U. Ionic liquids for aqueous two-phase extraction and stabilization of enzymes. *Biotechnol. Bioeng.* **2008**, *99*, 1416–1424.
14. Miloslavina, A.A.; Leipold, E.; Kijas, M.; Stark, A.; Heinemann, S.H.; Imhof, D. A room temperature ionic liquid as convenient solvent for the oxidative folding of conopeptides. *J. Pept. Sci.* **2009**, *15*, 72–77.
15. Plaquevent, J.C.; Levillain, J.; Guillen, F.; Malhiac, C.; Gaumont, A.C. Ionic liquids: New targets and media for alpha-amino acid and peptide chemistry. *Chem. Rev.* **2008**, *108*, 5035–5060.
16. Brown, R.; Carr, B.; Lennon, J. Factors that influence the observed fast fragmentation of peptides in matrix-assisted laser desorption. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 225–232.
17. McDonnell, L.A.; Heeren, R.M. Imaging mass spectrometry. *Mass Spectrom. Rev.* **2007**, *26*, 606–643.

18. Tietze, D.; Breitzke, H.; Imhof, D.; Kothe, E.; Weston, J.; Buntkowsky, G. New insight into the mode of action of nickel superoxide dismutase by investigating metallopeptide substrate models 1. *Chem. Eur. J.* **2009**, *15*, 517–523.
19. Becker, S.; Terlau, H. Toxins from cone snails: Properties, applications and biotechnological production. *Appl. Microbiol. Biotechnol.* **2008**, *79*, 1–9.
20. Escoubas, P.; Bernard, C.; Lambeau, G.; Lazdunski, M.; Darbon, H. Recombinant production and solution structure of PcTx1, the specific peptide inhibitor of ASIC1a proton-gated cation channels. *Protein Sci.* **2003**, *12*, 1332–1343.
21. Anangi, R.; Chen, C.Y.; Cheng, C.H.; Chen, Y.C.; Chen, C.C.; Chu, Y.P.; Chang, C.H.; Jeng, W.Y.; Shiu, J.H.; Chuang, W.J. Expression of snake venom toxins in *Pichia pastoris*. *Toxin Rev.* **2007**, *26*, 169–187.
22. Olivera, B.M.; Gray, W.R.; Zeikus, R.; McIntosh, J.M.; Varga, J.; Rivier, J.; de Santos, V.; Cruz, L.J. Peptide neurotoxins from fish-hunting cone snails. *Science* **1985**, *230*, 1338–1343.
23. Guzman, F.; Barberis, S.; Illanes, A. Peptide synthesis: Chemical or enzymatic. *J. Biotechnol.* **2007**, *10*, 279–314.
24. Marsh, K.N.; Deev, A.; Wu, A.C.T.; Tran, E.; Klamt, A. Room temperature ionic liquids as replacements for conventional solvents—A review. *Korean J. Chem. Eng.* **2002**, *19*, 357–362.
25. Kumar, D.; Bhalla, T.C. Microbial proteases in peptide synthesis: Approaches and applications. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 726–736.
26. Merrifield, R.B. Solid phase peptide synthesis 1. Synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
27. Mutter, M.; Hagenmai, H.; Bayer, E. New method of polypeptide synthesis. *Angew. Chem. Int. Ed.* **1971**, *10*, 811–812.
28. Horvath, I.T. Fluorous biphasic chemistry. *Acc. Chem. Res.* **1998**, *31*, 641–650.
29. Mizuno, M.; Goto, K.; Miura, T.; Hosaka, D.; Inazu, T. A novel peptide synthesis using fluororous chemistry. *Chem. Commun.* **2003**, 972–973.
30. Xing, G.W.; Li, F.Y.; Ming, C.; Ran, L.N. Peptide bond formation catalyzed by alpha-chymotrypsin in ionic liquids. *Tetrahedron Lett.* **2007**, *48*, 4271–4274.
31. Erbeltinger, M.; Mesiano, A.J.; Russell, A.J. Enzymatic catalysis of formation of Z-aspartame in ionic liquid—An alternative to enzymatic catalysis in organic solvents. *Biotechnol. Prog.* **2000**, *16*, 1129–1131.
32. Smith, V.F.; de Long, H.C.; Sutto, T.E.; Trulove, P.C. *Nonenzymatic Synthesis of Peptides in Ionic Liquids*; Proceedings of the Electrochemical Society, Molten Salts XIII, Annapolis, MD, USA, PV 2002-19; pp. 268–275.
33. Vallette, H.; Ferron, L.; Coquerel, G.; Gaumont, A.C.; Plaquevent, J.C. Peptide synthesis in room temperature ionic liquids. *Tetrahedron Lett.* **2004**, *45*, 1617–1619.
34. Vallette, H.; Ferron, L.; Coquerel, G.; Guillen, F.; Plaquevent, J.C. Room temperature ionic liquids (RTIL's) are convenient solvents for peptide synthesis. *ARKIVOC* **2006**, *iv*, 200–211.
35. Chan, T.-H.; He, X. Imidazolium-type Ionic Oligomers. U.S. Patent 2010/0093975 A1, April 15, 2010.
36. Miao, W.; Chan, T.H. Ionic-liquid-supported peptide synthesis demonstrated by the synthesis of Leu(5)-enkephalin. *J. Org. Chem.* **2005**, *70*, 3251–3255.

37. Kim, D.W.; Choe, Y.S.; Chi, D.Y. A new nucleophilic fluorine-18 labeling method for aliphatic mesylates: Reaction in ionic liquids shows tolerance for water. *Nucl. Med. Biol.* **2003**, *30*, 345–350.
38. Schirrmacher, R.; Wängler, C.; Schirrmacher, E. Recent developments and trends in ^{18}F -radiochemistry: Syntheses and applications. *Mini-Rev. Org. Chem.* **2007**, *4*, 317–329.
39. Miloslavina, A.; Ebert, C.; Tietze, D.; Ohlenschlager, O.; Englert, C.; Gorlach, M.; Imhof, D. An unusual peptide from *Conus villeginii*: Synthesis, solution structure, and cardioactivity. *Peptides* **2010**, *31*, 1292–1300.
40. Bihari, M.; Russell, T.P.; Hoagland, D.A. Dissolution and dissolved state of cytochrome c in a neat, hydrophilic ionic liquid. *Biomacromolecules* **2010**, *11*, 2944–2948.
41. Constantinescu, D.; Weingartner, H.; Herrmann, C. Protein denaturation by ionic liquids and the Hofmeister series: A case study of aqueous solutions of ribonuclease A. *Angew. Chem. Int. Ed.* **2007**, *46*, 8887–8889.
42. Noritomi, H.; Suzuki, K.; Kikuta, M.; Kato, S. Catalytic activity of α -chymotrypsin in enzymatic peptide synthesis in ionic liquids. *Biochem. Eng. J.* **2009**, *47*, 27–30.
43. Wang, Z.; Xiao, H.; Han, Y.; Jiang, P.; Zhou, Z. The effect of four imidazolium ionic liquids on hen egg white lysozyme solubility. *J. Chem. Eng. Data* **2011**, *56*, 1700–1703.
44. Summers, C.A.; Flowers, R.A., II. Protein renaturation by the liquid organic salt ethylammonium nitrate. *Protein Sci.* **2000**, *9*, 2001–2008.
45. Buchfink, R.; Tischer, A.; Patil, G.; Rudolph, R.; Lange, C. Ionic liquids as refolding additives: Variation of the anion. *J. Biotechnol.* **2010**, *150*, 64–72.
46. Sasmal, D.K.; Mondal, T.; Sen Mojumdar, S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. An FCS study of unfolding and refolding of CPM-labeled human serum albumin: Role of ionic liquid. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
47. Vasantha, T.; Attri, P.; Venkatesu, P.; Devi, R.S.R. Thermodynamic contributions of peptide backbone unit from water to biocompatible ionic liquids at $T = 298.15\text{ K}$. *J. Chem. Thermodyn.* **2012**, *45*, 122–136.
48. Wehofskey, N.; Wespe, C.; Cerovsky, V.; Pech, A.; Hoess, E.; Rudolph, R.; Bordusa, F. Ionic liquids and proteases: A clean alliance for semisynthesis. *ChemBioChem* **2008**, *9*, 1493–1499.
49. Ternois, J.; Ferron, L.; Coquerel, G.; Guillen, F.; Plaquevent, J.-C. Ionic Liquids: New Opportunities for the Chemistry of Amino Acids, Peptides, and Pharmaceutical Compounds. In *Ionic Liquid Applications: Pharmaceuticals, Therapeutics, and Biotechnology*; American Chemical Society, Oxford University Press: Cary, NC, USA, 2010; Volume 1038, pp. 13–24.
50. Malhotra, S.V.; Zhang, C.; Wang, H. Enzymatic synthesis of dipeptides in ionic liquids. *Lett. Org. Chem.* **2010**, *7*, 168–171.
51. Ferron, L.; Guillen, F.; Coste, S.; Coquerel, G.; Cardinaël, P.; Schwartz, J.; Paris, J.-M.; Plaquevent, J.-C. Design and scalable synthesis of new chiral selectors. Part 1: Synthesis and characterization of a new constrained cyclopeptide from unnatural bulky amino acids. *Tetrahedron* **2011**, *67*, 6036–6044.
52. Petiot, P.; Charnay, C.; Martinez, J.; Puttergill, L.; Galindo, F.; Lamaty, F.; Colacino, E. Synthesis of a new hydrophilic poly(ethylene glycol)-ionic liquid and its application in peptide synthesis. *Chem. Commun. (Camb)* **2010**, *46*, 8842–8844.

53. He, X.; Chan, T.H. Structurally defined imidazolium-type ionic oligomers as soluble/solid support for peptide synthesis. *Org. Lett.* **2007**, *9*, 2681–2684.
54. Bonnette, F.; Mincheva, Z.; Lavastre, O. Functionalized ionic liquids as new supports for peptide coupling and traceless catalyzed carbon-carbon coupling reactions. *Comb. Chem. High Throughput Screen.* **2006**, *9*, 229–232.
55. Mincheva, Z.; Bonnette, F.; Lavastre, O. Ionic liquid supports stable under conditions of peptide couplings, deprotections and traceless Suzuki reactions. *Collect. Czech. Chem. Commun.* **2007**, *72*, 417–434.
56. Guillen, F.; Brégeon, D.; Plaquevent, J.-C. (S)-Histidine: The ideal precursor for a novel family of chiral aminoacid and peptidic ionic liquids. *Tetrahedron Lett.* **2006**, *47*, 1245–1248.
57. Chen, L.; Zheng, M.F.; Zhou, Y.; Liu, H.; Jiang, H.L. Ionic-liquid-supported total synthesis of sansalvamide a peptide. *Synth. Commun.* **2008**, *38*, 239–248.
58. Roche, C.; Pucheault, M.; Vaultier, M.; Commerçon, A. Onium salt supported peptide synthesis. *Tetrahedron* **2010**, *66*, 8325–8334.
59. Cho, H.-J.; Lee, S.-M.; Jung, S.; Lee, T.-K.; Yoon, H.-J.; Lee, Y.-S. Ionic liquid incorporated polystyrene resin for solid-phase peptide synthesis. *Tetrahedron Lett.* **2011**, *52*, 1459–1461.
60. Bolton, R. Radiohalogen incorporation into organic systems. *J. Label. Compd. Radiopharm.* **2002**, *45*, 485–528.
61. Moroder, L.; Besse, D.; Musiol, H.J.; RudolphBohner, S.; Siedler, F. Oxidative folding of cystine-rich peptides vs. regioselective cysteine pairing strategies 1. *Biopolymers* **1996**, *40*, 207–234.
62. Moroder, L.; Musiol, H.J.; Gotz, M.; Renner, C. Synthesis of single- and multiple-stranded cystine-rich peptides. *Biopolymers* **2005**, *80*, 85–97.
63. Moroder, L.; Musiol, H.-J.; Schaschke, N.; Chen, L.; Hargittal, B.; Barany, G. *Methods of Organic Chemistry, Synthesis of Peptides and Peptidomimetics*; Thieme-Verlag, S., Goodman, M., Felix, A., Moroder, L., Toniolo, C., Eds.; Houben-Weyl, Georg Thieme Verlag: Stuttgart, Germany, 2003; Volume E22a, pp. 384–423.
64. Besser, D.; Muller, B.; Agricola, I.; Reissmann, S. Synthesis of differentially protected N-acylated reduced pseudodipeptides as building units for backbone cyclic peptides. *J. Pept. Sci.* **2000**, *6*, 130–138.
65. Reissmann, S.; Imhof, D. Development of conformationally restricted analogues of bradykinin and somatostatin using constrained amino acids and different types of cyclization. *Curr. Med. Chem.* **2004**, *11*, 2823–2844.
66. Buczek, P.; Buczek, O.; Bulaj, G. Total chemical synthesis and oxidative folding of delta-conotoxin PVIA containing an N-terminal propeptide. *Biopolymers* **2005**, *80*, 50–57.
67. Angell, Y.M.; Alsina, J.; Albericio, F.; Barany, G. Practical protocols for stepwise solid-phase synthesis of cysteine-containing peptides. *J. Pept. Res.* **2002**, *60*, 292–299.
68. Han, Y.; Albericio, F.; Barany, G. Occurrence and minimization of cysteine racemization during stepwise solid-phase peptide synthesis. *J. Org. Chem.* **1997**, *62*, 4307–4312.
69. DeLa Cruz, R.; Whitby, F.G.; Buczek, O.; Bulaj, G. Detergent-assisted oxidative folding of delta-conotoxins. *J. Pept. Res.* **2002**, *61*, 202–212.

70. Fuller, E.; Green, B.R.; Catlin, P.; Buczek, O.; Nielsen, J.S.; Olivera, B.M.; Bulaj, G. Oxidative folding of conotoxins sharing an identical disulfide bridging framework. *FEBS J.* **2005**, *272*, 1727–1738.
71. Lin Huang, J.; Noss, M.E.; Schmidt, K.M.; Murray, L.; Bunagan, M.R. The effect of neat ionic liquid on the folding of short peptides. *Chem. Commun.* **2011**, *47*, 8007–8009.
72. Debeljuh, N.; Barrow, C.J.; Henderson, L.; Byrne, N. Structure inducing ionic liquids-enhancement of alpha helicity in the Abeta(1–40) peptide from Alzheimer's disease. *Chem. Commun.* **2011**, *47*, 6371–6373.
73. Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779.
74. Kent, S.B. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.
75. Muir, T.W.; Dawson, P.E.; Kent, S.B. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* **1997**, *289*, 266–298.
76. Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H.U.; Lau, H. Über Peptidsynthesen. 8. Mitteilung Bildung von S-haltigen Peptiden durch intramolekulare Wanderung von Aminoacylresten. *Liebigs Ann. Chem.* **1953**, *583*, 129–149.
77. Romanelli, A.; Shekhtman, A.; Cowburn, D.; Muir, T.W. Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6397–6402.
78. Böhm, M.; Kühn, T.; Hards, K.; Coch, R.; Arkona, C.; Schlott, B.; Steinmetzer, T.; Imhof, D. Synthesis and functional characterization of tridegin and its analogues: Inhibitors and substrates of factor XIIIa. *ChemMedChem* **2012**, *7*, 326–333.
79. Hackeng, T.M.; Griffin, J.H.; Dawson, P.E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
80. Johnson, E.C.; Kent, S.B. Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
81. Dawson, P.E.; Churchill, M.J.; Ghadiri, M.R.; Kent, S.B. Modulation of reactivity in NCL through the use of thiol additives. *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
82. Escher, S.E.; Klüver, E.; Adermann, K. Fmoc-based synthesis of the human CC chemokine CCL14/HCC-1 by SPPS and native chemical ligation. *Lett. Pept. Sci.* **2002**, *8*, 349–357.
83. Han, D.; Row, K. Recent applications of ionic liquids in separation technology. *Molecules* **2010**, *15*, 2405–2426.
84. Polyakova, Y.; Koo, Y.; Row, K. Application of ionic liquids as mobile phase modifier in HPLC. *Biotechnol. Bioprocess Eng.* **2006**, *11*, 1–6.
85. Polyakova, Y.; Koo, Y.; Row, K. Application of ionic liquids of some bioactive molecules in RP-HPLC. *Rev. Anal. Chem.* **2007**, *26*, 77–98.
86. Wang, Y.; Tian, M.; Bi, W.; Row, K. Application of ionic liquids in high performance reversed-phase chromatography. *Int. J. Mol. Sci.* **2009**, *10*, 2591–2610.
87. Chitta, K.; Van Meter, D.; Stalcup, A. Separation of peptides by HPLC using a surface-confined ionic liquid stationary phase. *Anal. Bioanal. Chem.* **2010**, *396*, 775–781.

88. Marszall, M.; Kaliszan, R. Application of ionic liquids in liquid chromatography. *Crit. Rev. Anal. Chem.* **2007**, *37*, 127–140.
89. Abraham, M.; Poole, C.; Poole, S. Solute effects on reversed-phase thin-layer chromatography A linear free energy relationship analysis. *J. Chromatogr. A* **1996**, *749*, 201–209.
90. Abraham, M.; Treiner, C.; Roses, M.; Rafols, C.; Ishihama, Y. Linear free energy relationship analysis of microemulsion electrokinetic chromatographic determination of lipophilicity. *J. Chromatogr. A* **1996**, *752*, 243–249.
91. Van Meter, D.; Sun, Y.; Parker, K.; Stalcup, A. Retention characteristics of a new butylimidazolium-based stationary phase. Part II: Anion exchange and partitioning. *Anal. Bioanal. Chem.* **2008**, *390*, 897–905.
92. Qiu, H.; Jiang, S.; Liu, X. *N*-methylimidazolium anion-exchange stationary phase for high-performance liquid chromatography. *J. Chromatogr. A* **2006**, *1103*, 265–270.
93. Van Meter, D.; Oliver, N.; Carle, A.; Dehm, S.; Ridgway, T.; Stalcup, A. Characterization of surface-confined ionic liquid stationary phases: Impact of cation and anion identity on retention. *Anal. Bioanal. Chem.* **2009**, *393*, 283–294.
94. Kaliszan, R.; Marszall, M.; Markuszewski, M.; Baczek, T.; Pernak, J. Suppression of deleterious effects of free silanols in liquid chromatography by imidazolium tetrafluoroborate ionic liquids. *J. Chromatogr. A* **2004**, *1030*, 263–271.
95. Marszall, M.; Baczek, T.; Kaliszan, R. Reduction of silanophilic interactions in liquid chromatography with the use of ionic liquids. *Anal. Chim. Acta* **2005**, *547*, 172–178.
96. He, L.; Zhang, W.; Zhao, L.; Liu, X.; Jiang, S. Effect of 1-alkyl-3-methylimidazolium-based ionic liquids as the eluent on the separation of ephedrine by liquid chromatography. *J. Chromatogr. A* **2003**, *1007*, 39–45.
97. Marszall, M.; Baczek, T.; Kaliszan, R. Evaluation of the silanol-suppressing potency of ionic liquids. *J. Sep. Sci.* **2006**, *29*, 1138–1145.
98. Baczek, T.; Sparzak, B. Ionic liquids as novel solvent additives to separate peptides. *Z. Naturforsch C* **2006**, *61*, 827–832.
99. Baczek, T.; Marszall, M.P.; Kaliszan, R.; Walijewski, L.; Makowiecka, W.; Sparzak, B.; Grzonka, Z.; Wisniewska, K.; Juszczak, P. Behavior of peptides and computer-assisted optimization of peptides separations in a normal-phase thin-layer chromatography system with and without the addition of ionic liquid in the eluent. *Biomed. Chromatogr.* **2005**, *19*, 1–8.
100. Mowthorpe, S.; Clench, M.; Cricelius, A.; Richards, D.; Parr, V.; Tetler, L. Matrix-assisted laser desorption/ionisation time-of-flight thin layer chromatography mass spectrometry—A rapid method for impurity testing. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 264–270.
101. Todd, P.J.; Schaaff, T.G.; Chaurand, P.; Caprioli, R.M. Organic ion imaging of biological tissue with secondary ion mass spectrometry and matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* **2001**, *36*, 355–369.
102. Jackson, S.; Wang, H.; Woods, A. Direct tissue analysis of phospholipids in rat brain using MALDI-TOFMS and MALDI-ion mobility-TOFMS. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 133–138.
103. Jones, J.J.; Batoy, S.M.; Wilkins, C.L.; Liyanage, R.; Lay, J.O., Jr. Ionic liquid matrix-induced metastable decay of peptides and oligonucleotides and stabilization of phospholipids in MALDI FTMS analyses. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 2000–2008.

104. Towers, M.W.; McKendrick, J.E.; Cramer, R. Introduction of 4-chloro- α -cyanocinnamic acid liquid matrices for high sensitivity UV-MALDI MS. *J. Proteome Res.* **2010**, *9*, 1931–1940.
105. Zabet-Moghaddam, M.; Kruger, R.; Heinzle, E.; Tholey, A. Matrix-assisted laser desorption/ionization mass spectrometry for the characterization of ionic liquids and the analysis of amino acids, peptides and proteins in ionic liquids. *J. Mass Spectrom.* **2004**, *39*, 1494–1505.
106. Mank, M.; Stahl, B.; Boehm, G. 2,5-Dihydroxybenzoic acid butylamine and other ionic liquid matrixes for enhanced MALDI-MS analysis of biomolecules. *Anal. Chem.* **2004**, *76*, 2938–2950.
107. Tholey, A.; Heinzle, E. Ionic (liquid) matrices for matrix-assisted laser desorption/ionization mass spectrometry-applications and perspectives. *Anal. Bioanal. Chem.* **2006**, *386*, 24–37.
108. Zabet-Moghaddam, M.; Heinzle, E.; Lasasosa, M.; Tholey, A. Pyridinium-based ionic liquid matrices can improve the identification of proteins by peptide mass-fingerprint analysis with matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Bioanal. Chem.* **2006**, *384*, 215–224.
109. Tholey, A.; Zabet-Moghaddam, M.; Heinzle, E. Quantification of peptides for the monitoring of protease-catalyzed reactions by matrix-assisted laser desorption/ionization mass spectrometry using ionic liquid matrixes. *Anal. Chem.* **2006**, *78*, 291–297.
110. Tholey, A. Ionic liquid matrices with phosphoric acid as matrix additive for the facilitated analysis of phosphopeptides by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1761–1768.
111. Palmblad, M.; Cramer, R. Liquid matrix deposition on conductive hydrophobic surfaces for tuning and quantitation in UV-MALDI mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 693–697.
112. Hortal, A.; Hurtado, P.; Martinez-Haya, B. Matrix-assisted laser desorption mass spectrometry of gas-phase peptide-metal complexes. *Appl. Phys. A: Mater. Sci. Process.* **2008**, *93*, 935–939.
113. Calvano, C.; Carulli, S.; Palmisano, F. Aniline/ α -cyano-4-hydroxycinnamic acid is a highly versatile ionic liquid for matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 1659–1668.
114. Crank, J.; Armstrong, D. Towards a second generation of ionic liquid matrices (ILMs) for MALDI-MS of peptides, proteins, and carbohydrates. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1790–1800.
115. Ullmer, R.; Rizzi, A. Use of a novel ionic liquid matrix for MALDI-MS analysis of glycopeptides and glycans out of total tryptic digests. *J. Mass Spectrom.* **2009**, *44*, 1596–1603.
116. Giménez, E.; Benavente, F.; Barbosa, J.; Sanz-Nebot, V. Ionic liquid matrices for MALDI-TOF-MS analysis of intact glycoproteins. *Anal. Bioanal. Chem.* **2010**, *398*, 357–365.
117. Fitzgerald, J.J.; Kunnath, P.; Walker, A.V. Matrix-enhanced secondary ion mass spectrometry (ME SIMS) using room temperature ionic liquid matrices. *Anal. Chem.* **2010**, *82*, 4413–4419.
118. Sun, L.; Tao, D.; Han, B.; Ma, J.; Zhu, G.; Liang, Z.; Shan, Y.; Zhang, L.; Zhang, Y. Ionic liquid 1-butyl-3-methyl imidazolium tetrafluoroborate for shotgun membrane proteomics. *Anal. Bioanal. Chem.* **2011**, *399*, 3387–3397.

Epilogue

This review focuses on recent application and progress of ILs in peptides synthesis. Although, the large number of different ILs with specific properties can be confusing and complicated, the review facilitates the decision for a specific IL in a distinct application. IL-supported peptide synthesis as well as chemical and enzymatic peptide synthesis in ILs represent novel applications of ILs based on aspects such as a higher solubility of poorly soluble educts and an increased product yield compared to the use of common solvents. Furthermore, oxidative folding, native chemical ligation and radiolabelling of peptides in ILs were included in peptide chemistry. In addition, immobilized ILs can be used for purification using thin-layer chromatography, RP-HPLC or as additives in the mobile phase. For analytical purposes, the negligible vapour pressure of ILs enables their use as a suitable matrix for MALDI-MS. Although the use of ILs in peptide chemistry has just started, the ongoing research interest and new findings open up much broader applications within distinct research areas. Studies supporting these developments are described in parts II and III.

4.2. Ionic Liquids as reaction media for oxidative folding and native chemical ligation of cysteine-containing peptides

Full Paper

Authors

Miriam Böhm, Alesia A. Tietze, Pascal Heimer, Ming Chen, Diana Imhof

Journal of Molecular Liquids, Volume 192, 2014, Pages 67-70.

doi: 10.1016/j.molliq.2013.08.020

Copyright © 2013 Elsevier B.V. All rights reserved.

Preface

After a comprehensive summary of the already published results comprising Ionic Liquids as alternative reaction media in peptide chemistry in the previous review, we intended to extend the application of ILs to oxidative folding and native chemical ligation of bioactive, cysteine-containing peptides.²³⁴ The anticoagulant 66mer tridegin and a model peptide LYRAXCRANK (X = any amino acid) were subjected to NCL in different solvents including ILs.^{235,236} For oxidative folding several conotoxins such as CCAP-*vil*, μ -SIIIA, μ -PIIIA, δ -EVIA and δ -SVIE with up to three disulfide bridges were subjected to disulfide bond formation in ILs. Furthermore, different factors influencing the reaction outcome were analyzed, including the use of different ILs, different temperatures and varying water content, in order to increase our knowledge in this respect. A more detailed analysis employing μ -SIIIA as model peptide is described in part III.



Ionic liquids as reaction media for oxidative folding and native chemical ligation of cysteine-containing peptides



Miriam Böhm, Alesia A. Tietze, Pascal Heimer, Ming Chen, Diana Imhof*

Pharmaceutical Chemistry I, Institute of Pharmacy, University of Bonn, Brühler Str. 7, D-53119 Bonn, Germany

ARTICLE INFO

Available online 13 September 2013

Keywords:

Ionic liquids
Native chemical ligation
Oxidative folding
Peptides
Cysteine-rich

ABSTRACT

This paper represents the final report on the results obtained during a two-years funding period within the DFG priority program 1191: Ionic liquids. The studies were focused on the interaction of ionic liquids with peptides in two distinct reactions involving cysteine residues, namely oxidative folding and native chemical ligation (NCL). Both reactions are generally performed in solution phase. With the exception of selected model compounds, the bioactive peptides used herein are rather hydrophobic and long (>20 amino acids) representatives of their respective class, e.g. conotoxins and enzyme inhibitors. Thus, ionic liquids were used as alternative media to circumvent solubility issues. We found that imidazolium-based ionic liquids, and more precisely [C₂mim][OAc], were most suitable for both reactions being superior to conventional aqueous buffer systems in several aspects, including reaction yield and time, concentration of starting material and product as well as selectivity.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The unique properties of ionic liquids (ILs) as a new class of solvents have been discussed extensively in the last decade [1–3]. Among the most interesting of these properties are the safety towards both, the chemist and the environment, including non-flammability, negligible vapor pressure and non-toxicity. Additionally, ILs are outstanding in their chemical variability, providing the chemical community with tools to tailor the solvent of their needs. However, the link between distinct changes in the IL composition and the applicability for a defined reaction still needs to be investigated in more detail to exploit the full potential of tailor-made solvents, co-solvents and additives.

Being rather restricted in the choice of solvent (primarily aqueous buffer systems similar to the physiological environment), biochemists have yet shown increasing interest in the application of ILs as solvents and/or co-solvents for different classes of biomolecules. Recent reports demonstrated the suitability of ILs as solvents and reaction media for e.g. carbohydrates [4,5] and proteins [6,7], in particular enzymes [3,8]. Ionic liquids seemed to be suitable to stabilize biological macromolecules in solution, and thus, also attracted attention for other applications such as those occurring in peptide chemistry, including

synthesis, purification/extraction, analytics and structure elucidation [9]. The aim of our studies was therefore to use ILs for selected reactions employing peptides, with a focus on cysteine-rich bioactive representatives of native origin. These peptides are medium-sized to long biomolecules (9-mer to 66-mer) and contain two to six cysteine residues in their sequence. There are two main reasons why the behavior of cysteine residues in ionic liquids is of interest to us: First, cysteine residues are required for native chemical ligation (NCL) of unprotected peptide fragments (Fig. 1) [10,11], and second, they play an important role in peptide structure formation due to the generation of disulfide bonds (oxidative folding) [12]. Both reactions can be carried out successfully in aqueous reaction media [11,12]. However, there are some drawbacks and difficulties accompanying oxidative folding and NCL when performed in buffer systems.

Oxidation of cysteine residues, i.e. the formation of disulfide bonds in peptides, is a major step in the synthesis of many biologically active, cysteine-containing peptides, such as cystine-knot peptides/conopeptides, cardioactive peptides, peptide hormones and others. In most cases, the oxidation step is carried out in aqueous buffer solutions in the presence of redox active agents (e.g. oxidized and reduced glutathione) exploiting the self-folding capability of the peptide's amino acid sequence rather than applying complex protecting group strategies [12].

Native chemical ligation was first introduced by Kent and co-workers in 1994 [10]. It provided an easy way to connect two unprotected peptide fragments resulting in a native peptide bond at the ligation site. However, as the method is based on a transthioesterification (Fig. 1), there is an absolute requirement for a cysteine residue on the N-terminus of the C-terminal fragment. In addition, the N-terminal segment should terminate as a thioester. Further investigation by Hackeng

Abbreviations: [OAc][−], acetate; [C₄mim]⁺, 1-butyl-3-methylimidazolium; [N(CN)₂][−], dicyanamide; [DEP][−], diethylphosphate; [DHP], dihydrogenphosphate; EAF, ethylammonium formate; [C₂mim]⁺, 1-ethyl-3-methylimidazolium; HPLC, high performance liquid chromatography; IL, ionic liquid; MS, mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; MAF, methylammonium formate; NCL, native chemical ligation; RTIL, room temperature ionic liquid; TFA, trifluoroacetic acid; TOF, time of flight; [OTs][−], tosylate.

* Corresponding author. Tel.: +49 228 7360258; fax: +49 228 7360223.

E-mail address: dimhof@uni-bonn.de (D. Imhof).

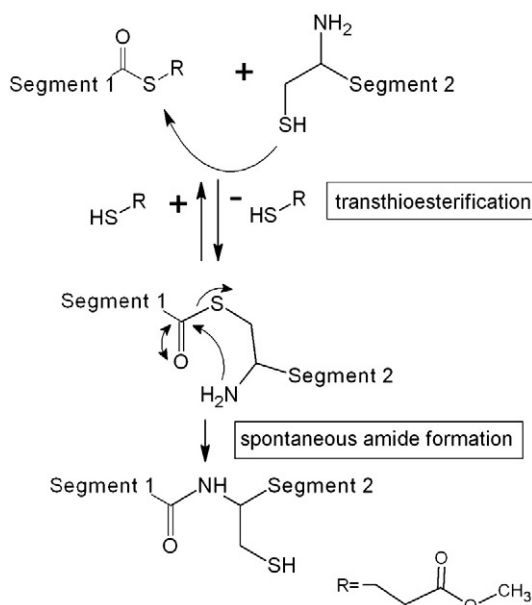


Fig. 1. Principle and mechanism of native chemical ligation [10].

et al. showed that the reaction rate and outcome seriously depend on the C-terminal amino acid of the N-terminal fragment [11]. Previous investigations in our group had shown an improvement of cysteine oxidation in distinct ILs [12,13]. Therefore, we decided to also use ILs as alternative media for NCL. After first promising results on the application of ILs in both reaction types, we started a more systematic investigation on the parameters influencing the reaction process and outcome, including IL cation and anion choice, temperature, water content and peptide choice, which is still ongoing. The first results of these efforts are summarized in this report.

2. Experimental

Peptide synthesis reagents and building blocks were purchased from Orpegen Peptide Chemicals GmbH (Heidelberg, Germany) and Iris Biotech GmbH (Marktredwitz, German), respectively. Ionic liquids were either kindly provided by Dr. Annegret Stark (University of Leipzig, Germany) or purchased from IoLiTec - Ionic Liquids Technologies GmbH (Heilbronn, Germany).

2.1. Synthesis

Peptide synthesis was performed automatically according to a standard Fmoc (N-(9-fluorenyl)-methoxycarbonyl) protocol, as earlier described [12–14].

2.2. Purification and characterization

Peptide purification after synthesis was performed using a semi-preparative HPLC instrument equipped with a C18 column (Knauer Eurospher 100) and applying a gradient elution system with eluent A being 0.1% TFA in water, and eluent B 0.1% TFA in acetonitrile/water (90:10). The purity of the products before and after purification was determined by analytical RP-HPLC with a gradient elution using the aforementioned eluent A and 0.1% TFA in acetonitrile as eluent B. Peptide identity was confirmed by MALDI-TOF mass spectrometry on a MALDI-TOF/TOF Ultraflex I Bruker mass spectrometer or an autoflex II TOF/TOF mass spectrometer. In both cases, α -cyano-4-hydroxycinnamic acid was used as matrix.

2.3. Oxidative folding

Peptides were subjected to oxidation in aqueous solutions as reported elsewhere [12]. For oxidative folding in ILs, the linear peptide was dissolved in a small volume of the corresponding IL (40–800 μ l) at a concentration between 3 and 20 mM and left at room temperature with gentle stirring. Work up was performed as previously described [9].

2.4. Native chemical ligation

NCL in buffer solutions was carried out according to literature reports by Dawson *et al.* and Hackeng *et al.* [10,11]. The procedure for native chemical ligation in ILs was carried out as reported in Böhm *et al.* [14].

3. Results and discussion

3.1. Oxidative folding

As described in [12], the oxidation of various cysteine-rich peptides in the ionic liquid [C₂mim][OAc] revealed very promising results (Fig. 2). Though [C₂mim][OAc] already turned out best in a selection of e.g. imidazolium-based ionic liquids with varying anions, our aim was to investigate peptide oxidation in further ionic liquids which recently showed great potential with regard to proteins and enzymes, respectively [3,6,7]. All of the peptides we used are bioactive compounds [12–14], and thus, beside the improvement of synthetic methods we were interested in maintaining bioactivity after processing in ILs, too. Nevertheless, in a two-years funding period work had to be structured and, thus, model peptides were selected first to investigate the change of different parameters for the respective reaction. These changes include the variation of IL cations and anions, water content, reaction temperature and time, concentration of peptide and/or IL as well as practical issues such as technical performance.

As the first and most simple model peptide, CCAP-vil was chosen, which is a cardioactive peptide first isolated from *Conus vilipinii* [13]. This compound is not strictly a cysteine-rich peptide, since it contains only one disulfide bond. However, due to the rather simple nature of its oxidation process and the reportedly high yield for oxidation in aqueous systems, it proved a valuable tool to rapidly test a number of ILs with respect to their general suitability for the approach [13]. The most suitable ionic liquid for the oxidation of CCAP-vil was [C₂mim][OTf] (3% H₂O, v/v), with a yield of more than 90% although carried out at room temperature. Furthermore, the reaction was carried out with increased water content (3, 5, 7 and 10%) and at higher temperatures (40, 60 and 80 °C). Whereas higher water concentrations turned out less favorable for our reaction, the temperature increase accelerated the reaction (unpublished results). The same effect was found for other peptides, e.g. μ -SIIIA, as well. The availability of solution NMR structural information of CCAP-vil [13] allows for molecular dynamics simulations in different media, which are currently undertaken to provide a more scientific explanation of the effects observed in the experiments and to reveal insights into the molecular basis of our observations of peptide-IL interactions.

The second model peptide used for oxidative folding in ILs was the μ -conotoxin SIIIA [12]. This peptide, in contrast to CCAP-vil, contains three disulfide bonds, and therefore has a much higher potential to form different isomers compared to CCAP-vil. However, it can still easily be oxidized in buffer systems, yielding only one major compound (Fig. 2). More detailed experiments with this peptide showed, as already seen before, that the most useful IL for this application was [C₂mim][OAc], while [C₄mim][OAc] still gave the desired reaction product, albeit with a lower yield [12]. The other anions tested previously for folding of μ -SIIIA, namely [OTf][−], [DEP][−] and [N(CN)₂][−], gave no reaction product at all [9,12]. With these findings in mind, further investigations were focused on ILs with H-bonding abilities, such as the protic ILs ethylammonium formate (EAF) and methylammonium formate (MAF). Both

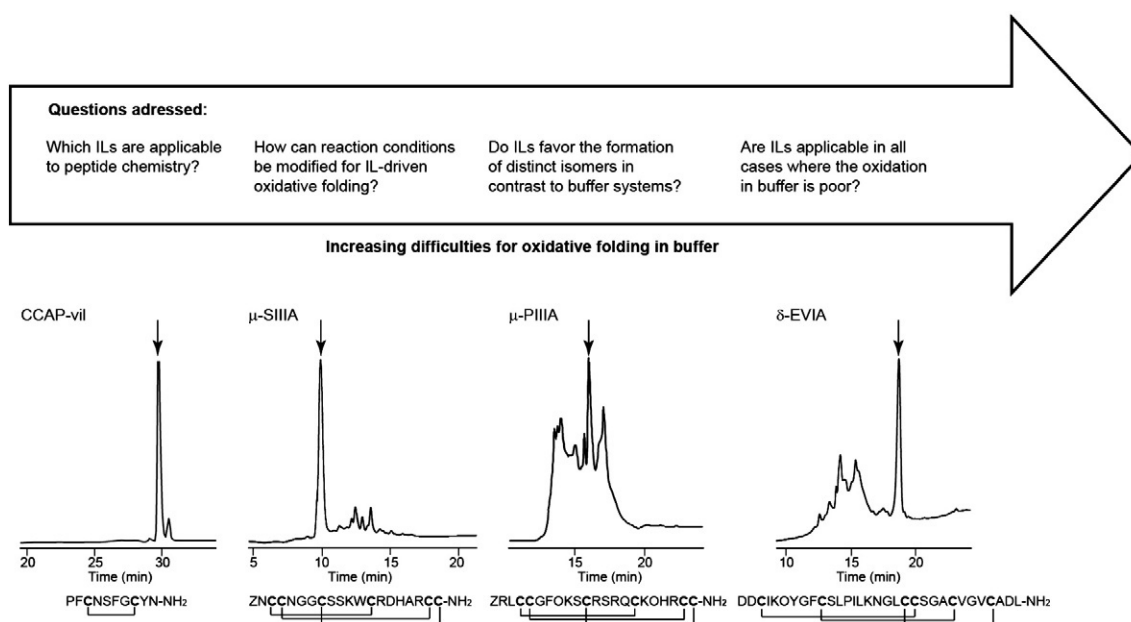


Fig. 2. General approach highlighting the selection process: moving from simple to more demanding model peptides with the aim to elucidate general aspects of IL-peptide-interaction. The HPLC profiles show crude products after oxidative folding in buffer, arrows indicate the proposed “native fold”, as shown below the profiles. (Z, pyroglutamic acid; O, hydroxyproline).

gave the desired product in relatively high yields (64% and 43%, respectively, [15] in contrast to other ILs (e.g. choline dihydrogenphosphate, 2 M solution), which did not result in the desired product. Therefore, studies will be repeated and continued to obtain a complete picture of whether and which ILs are applicable for peptide oxidation.

A more difficult peptide with respect to its self-folding ability in aqueous systems is another μ -conotoxin, namely μ -PIIIA. Oxidation in buffer resulted in a number of different completely and incompletely oxidized isomers of the reduced precursor, some of which have been investigated in more detail [16]. The occurrence of these isomers leads to a significant reduction of the expected version possessing the so-called “native fold” (as found in case of μ -SIIIA). Surprisingly, only one major product is formed if the reaction is performed in [C₂mim][OAc] [12]. Current experiments aim at the extraction of this product from the IL reaction mixture in order to provide sufficient amount for structural investigations as done for the buffer-oxidized isomers [16]. Simultaneously, first experiments have been performed with the even more hydrophobic δ -conotoxin EVIA and SVIE [12]. Unfortunately, these peptides are very difficult to handle and thus, more efforts are required to obtain their NMR solution structure from the products of the oxidative folding approach in both media. In buffer, the desired disulfide-bridged peptides are only obtained in very low yields (<10%), yet the results with [C₂mim][OAc] [12] were again promising for our future studies with these conotoxins.

3.2. Native chemical ligation

First experiments with regard to native chemical ligation in ionic liquids were performed using the anticoagulant peptide tridegin as subject of analysis [14]. In contrast to the reaction in a buffer system, where additives are required to succeed, the reaction in [C₂mim][OAc] proceeded smoothly even in the absence of any additive and at room temperature. The yield of 45% was even higher than in buffer containing thiophenol and benzylmercaptan (19%). Encouraged by these results, we used the funding period to start with investigations concerning NCL with model peptides. An approach focusing on the impact of the C-terminal amino acid of the N-terminal ligation segment, as suggested earlier by Hackeng et al. [11] was initiated with the sequence

LYRAXCRANK (X = any amino acid). The different LYRAX-peptide thioesters and the C-terminal CRANK-sequence were synthesized on the solid phase and subjected to ligation in ILs, primarily [C₂mim][OAc]. First results with the peptide LYRAN as the N-terminal fragment revealed that the ligation occurs much faster than in buffer (Fig. 3). Further studies are in progress with other amino acids (e.g. X = Gly, Leu, Val, and Lys) to provide information for a more generally applicable protocol.

4. Conclusions

4.1. Oxidative folding

From the findings on oxidative folding of peptides in ILs, different conclusions can be drawn, the most important is that there are numerous parameters influencing the quality and yield of the product [12–16]. First, the peptide itself, e.g. primary sequence, length, and net charge, determines the success of the reaction performed in an IL. Second, the IL itself has a great impact on the oxidation process, not only because of the variability in cations and anions, but also due to aspects such as water content. Indeed, the latter was shown to have a detrimental effect on the oxidation process in all cases tested that is in contrast to IL applications in protein research where in the majority of studies water was highly essential to maintain folding and activity [3]. A temperature increase, on the other hand, e.g. to 60 °C or even 80 °C seems to increase reaction rate without damaging the dissolved molecules. This is consistent with findings for proteins which are also thermo-stabilized in ILs. From the current state of results we can conclude the following: i) The high concentrations of a peptide, especially hydrophobic ones, that can easily be dissolved in distinct ionic liquids, such as imidazolium-based ILs, compared to aqueous systems show that ILs are generally useful media for this type of reaction. Furthermore, despite a high peptide concentration no dimer or multimer formation was observed so far. This can be explained by the integration of individual molecules into the IL network as earlier suggested for enzymes in ILs [9]. Moreover, oligomerization (intermolecular reaction) is disfavored versus the intramolecular disulfide formation due to proximity of the thiol functions after prefolding of the linear sequence. Stability and aggregation issues

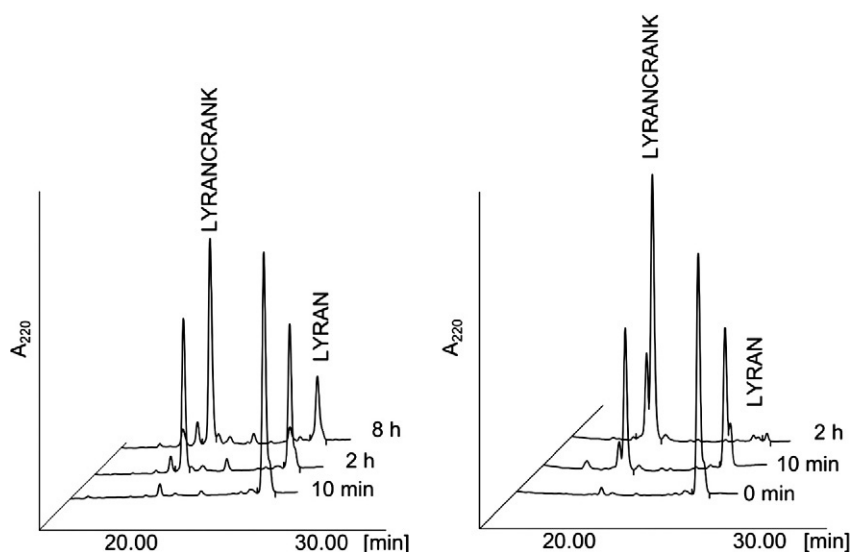


Fig. 3. HPLC profiles of the reaction mixture for the ligation reaction of LYRAN to CRANK-NH₂ performed at room temperature are shown. A) NCL in buffer is not finished after 8 h, while B) the reaction is completed after 2 h in [C₂mim][OAc].

were intensively studied for proteins in ILs, however, beside the high thermal stability of the peptides in ILs, differences between peptides and proteins with respect to the accepted water content seem to be caused by differences in the respective interactions with the IL cations and anions. Though similarities to proteins with respect to observed trends of a correlation with the Hofmeister ion series exist [3], peptides need to be discussed independently and need much more attention to catch up with the knowledge available for proteins.

ii) Up to now, the selection of an IL for peptide reaction, in particular oxidation, followed more or less a “trial-and-error” approach. However, it was found that the composition of the IL plays a major role with respect to reaction outcome. Generally, ILs that are not liquid at room temperature, but need the addition of water are not suitable for the applications described herein, as water had a negative influence on the reaction outcome. The tendency of the anion to accept hydrogen bonds seems to be preferable for highly charged peptides, whereas poorly charged representatives seem to require anions with a lower tendency to hydrogen bonding. Nevertheless, further investigations with e.g. negatively charged peptides are required to underpin this conclusion by extended and more systematic studies. Earlier investigations focused on imidazolium-based ILs, however, the herein presented first positive results obtained for selected protic ILs are encouraging with regard to further studies aiming at the definition of the requirements for a tailor-made IL for a specific reaction involving peptides.

4.2. Native chemical ligation

The same as indicated above is true for the second type of reaction studied within the two-years funding period. Preliminary results with both the bioactive 66-mer peptide tridegin [14], and the 10-mer model peptides again demonstrated the high potential of ionic liquids as reaction media. Reaction rate was much higher compared to the same reaction performed in the typical aqueous systems, and again, the concentration of the respective peptide fragments can be increased significantly due to a much better solubility in ILs compared to water-based media. The transthioesterification (intermolecular reaction) of two high concentrated reactants occur fast and smoothly suppressing

transformation of the thioester to the carboxylic acid and is favored because of the activated thioester and low water content. These first findings for NCL allow to speculate that the same effects observed for oxidative folding concerning water content and temperature can be expected here, and will be in the focus of our perspective investigations.

Acknowledgments

The authors like to thank the DFG for financial support within priority program SPP1911 “Ionic liquids”. For support and fruitful collaboration we are also grateful to the other members of the priority program, in particular A. Stark (University of Leipzig). We thank M. Engeser (University of Bonn) for access to the mass spectrometer facilities.

References

- [1] P. Wasserscheid, A. Stark, in: P.T. Anastas (Ed.) *Green Solvents*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim: Yale University, Center of Green Chemistry & Green Engineering, 2010, p. 39.
- [2] K.N. Marsh, A. Deev, A.C.T. Wu, E. Tran, A. Klamt, *Korean J. Chem. Eng.* 19 (2002) 357.
- [3] P. Domínguez de María, *Ionic Liquids in Biotransformations and Organocatalysis: Solvents and Beyond*, 1st ed. John Wiley & Sons, 2012.
- [4] R.C. Remsing, G. Hernandez, R.P. Swatloski, W.W. Masseski, R.D. Rogers, G. Moyna, *J. Phys. Chem. B* 112 (2008) 11071.
- [5] M. Sellin, B. Ondruschka, A. Stark, in: T.H.T. Liebert, K. Edgar (Eds.), *ACS Symp. Series*, Washington DC, 2010, pp. 121–135.
- [6] X. Chen, J. Liu, J. Wang, *Anal. Methods* 2 (2010) 1222.
- [7] H. Weingaertner, C. Cabrele, C. Herrmann, *Phys. Chem. Chem. Phys.* 14 (2012) 415.
- [8] M. Naushad, O.Z.A. Al, A.B. Khan, M. Ali, *Int. J. Biol. Macromol.* 51 (2012) 555.
- [9] A.A. Tietze, P. Heimer, A. Stark, D. Imhof, *Molecules* 17 (2012) 4158.
- [10] P.E. Dawson, T.W. Muir, I. Clark-Lewis, S.B. Kent, *Science* 266 (1994) 776.
- [11] T.M. Hackeng, J.H. Griffin, P.E. Dawson, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 10068.
- [12] A.A. Miloslavina, E. Leipold, M. Kijas, A. Stark, S.H. Heinemann, D. Imhof, *J. Pept. Sci.* 15 (2009) 72.
- [13] A. Miloslavina, C. Ebert, D. Tietze, O. Ohlenschläger, C. Englert, M. Gorch, D. Imhof, *Peptides* 31 (2010) 1292.
- [14] M. Böhm, T. Kühl, K. Hards, R. Coch, C. Arkona, B. Schlott, T. Steinmetzer, D. Imhof, *ChemMedChem* 7 (2012) 326.
- [15] P. Heimer, A.A. Tietze, M. Böhm, R. Giernoth, A. Kuchenbuch, A. Stark, E. Leipold, S.H. Heinemann, D. Imhof, *Chem. Eur. J.* (submitted for publication).
- [16] A.A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Kühl, A. Mischo, G. Buntkowsky, M. Gorch, S.H. Heinemann, D. Imhof, *Angew. Chem. Int. Ed. Engl.* 51 (2012) 4058.

Epilogue

This publication summarizes the results from a two-year study and the outcome described in three publications. First, results regarding the NCL reaction applied to model peptide LYRAXCRANK obtained in ILs provide that a faster reaction might occur in the IL [C₂mim][OAc] compared to the use of conventional buffer systems. Second, for oxidative folding carried out in ILs several parameters were investigated indicating major influence of the IL composition on product formation. Furthermore, high water content disfavoured, while an increase in temperature favoured product formation. Both, hydrophobic and hydrophilic conotoxins were dissolved in much higher concentrations in the respective IL compared to conventional oxidation in buffer without di- or oligomerization. This represents an advantage compared to common oxidation protocols using oxidizing agents, in particular for poorly soluble hydrophobic peptides, where high dilution is required to avoid dimer formation. Selected studies were completed meanwhile, which revealed further insight in the use of ILs in NCL (Kühl *et al.*)²³⁷ and oxidative folding (Heimer *et al.*). The latter publication is outlined in part III.

4.3 Application of room-temperature aprotic and protic Ionic Liquids for oxidative folding of cysteine-rich peptides

Full Paper

Authors

Pascal Heimer, Alesia A. Tietze, Miriam Böhm, Ralf Giernoth, Andrea Kuchenbuch, Annegret Stark, Enrico Leipold, Stefan H. Heinemann, Christian Kandt, Diana Imhof

ChemBioChem. 2014, 15, 18, 2754-65.

doi: 10.1002/cbic.201402356.

Copyright © 1999 - 2016 John Wiley & Sons, Inc. All Rights Reserved

Preface

The impact of IL composition (choice of cation and anion) in combination with several reaction parameters (educt concentration, temperature, water content) influences the product formation during oxidative folding of cysteine-rich peptides.²³⁴ This was investigated in order to expand the basic knowledge of this particular folding process carried out in suitable ILs. Herein, the conotoxin μ -SIIIA was used as a model peptide for the evaluation of reaction conditions in comparison with previously described results obtained in [C₂mim][OAc].¹⁶⁶ We intended to identify the most promising IL from a set of imidazolium-based, protic and aprotic ILs with respect to product formation. For the first time, bioactivity of μ -SIIIA oxidized in [C₂mim][OAc] was measured in electrophysiological studies using Na_v1.4-channel as the target protein. In addition, the effect of cation and anion of the ILs was investigated for the folding process of four conotoxins, namely α -GI, μ -SIIIA, μ -PIIIA, δ -EVIA, containing two or three cysteine bridges. This work represents an extension of the previous studies, i.e. Tietze et al. (2012) and Böhm et al. (2014) described in parts I and II.

Application of Room-Temperature Aprotic and Protic Ionic Liquids for Oxidative Folding of Cysteine-Rich Peptides

Pascal Heimer,^[a] Alesia A. Tietze,^[b] Miriam Böhm,^[a] Ralf Giernoth,^[c] Andrea Kuchenbuch,^[c] Annegret Stark,^[d] Enrico Leipold,^[e] Stefan H. Heinemann,^[e] Christian Kandt,^[f] and Diana Imhof^{*[a]}

The oxidation of the conotoxin μ -SIIIA in different ionic liquids was investigated, and the results were compared with those obtained in [C₂mim][OAc]. Conversion of the reduced precursor into the oxidized product was observed in the protic ILs methyl- and ethylammonium formate (MAF and EAF, respectively), whereas choline dihydrogenphosphate and Ammoeng 110 failed to yield folded peptide. However, the quality and yield of the peptide obtained in MAF and EAF were lower than in the case of the product from [C₂mim][OAc]. Reaction condi-

tions (temperature, water content) also had an impact on peptide conversion. A closer look at the activities of μ -SIIIA versions derived from an up-scaled synthesis in [C₂mim][OAc] revealed a significant loss of the effect on ion channel Na_v1.4 relative to the buffer-oxidized peptide, whereas digestion of either μ -SIIIA product by trypsin was unaffected. This was attributed to adherence of ions from the IL to the peptide, because the disulfide connectivity is basically the same for the differentially oxidized μ -SIIIA versions.

Introduction

Room-temperature ionic liquids (RTILs) were designed as "green solvents" to replace conventional volatile solvents with the aim of reducing their possible impacts on the environment, including ozone depletion, photochemical smog, and global climate change.^[1,2] The potential of RTILs has been extensively exploited over the last decade in fields as diverse as synthesis,^[3] analytical chemistry,^[4] electrochemistry,^[5] catalysis/biocatalysis,^[6–10] and materials science,^[11] at both laboratory and industrial scales.^[12,13] In addition, achievements for indus-

trial-scale production of "druggable" molecules are also highly important and still the subject of ongoing research activities, due to their potential for commercial applications in therapy and diagnostics. As a consequence, bioactive natural products have come more and more into the focus of applications of ionic liquids for their synthetic preparation starting from relatively inexpensive materials.

Recently we examined the use of selected RTILs for the preparation of multiply disulfide-bridged peptides and oligopeptides.^[4,14–16] The aim of thoroughly investigating bioactivity, mechanism of action, and therapeutic potential of such cysteine-rich peptides from snakes, scorpions, marine cone snails, leeches, or plants is still hampered by their inefficient and partially incorrect synthesis, accompanied by the formation of rather large amounts of misfolded by-products. This leads to limited availability, combined with insufficient quantities and purities required for detailed investigations. Numerous strategies, ranging from self-assembly to immobilization of folding reagents and/or components on solid supports, demonstrate the need to improve methods for oxidative refolding of disulfide-rich peptides and proteins.^[17,18] The concept of "integrated oxidative folding", which combines the use of diselenide and selectively (¹⁵N/¹³C)-labeled disulfide bridges, has been introduced in order to achieve both improved folding yields and simultaneous identification of correctly folded species by NMR spectroscopy.^[17,18] This approach, however, requires the incorporation of suitably protected selenocysteine units instead of the naturally occurring cysteine residues.

The first objective of the current study was therefore to examine whether application of RTILs for reactions involving cysteine might lead to improvements in accessibility and yields of naturally occurring Cys-rich peptides, on the basis that ILs

[a] P. Heimer,[†] M. Böhm,[†] Prof. Dr. D. Imhof
Department of Pharmaceutical Chemistry I
Pharmaceutical Institute, University of Bonn
Brühler Strasse 7, 53119 Bonn (Germany)
E-mail: dimhof@uni-bonn.de

[b] Dr. A. A. Tietze[†]
Technische Universität Darmstadt
Clemens-Schöpf Institute of Organic and Biochemistry
Alarich-Weiss-Strasse 4, 64287 Darmstadt (Germany)


[c] Dr. R. Giernoth, A. Kuchenbuch
Department of Chemistry, University of Cologne
Greinstrasse 4, 50939 Köln (Germany)

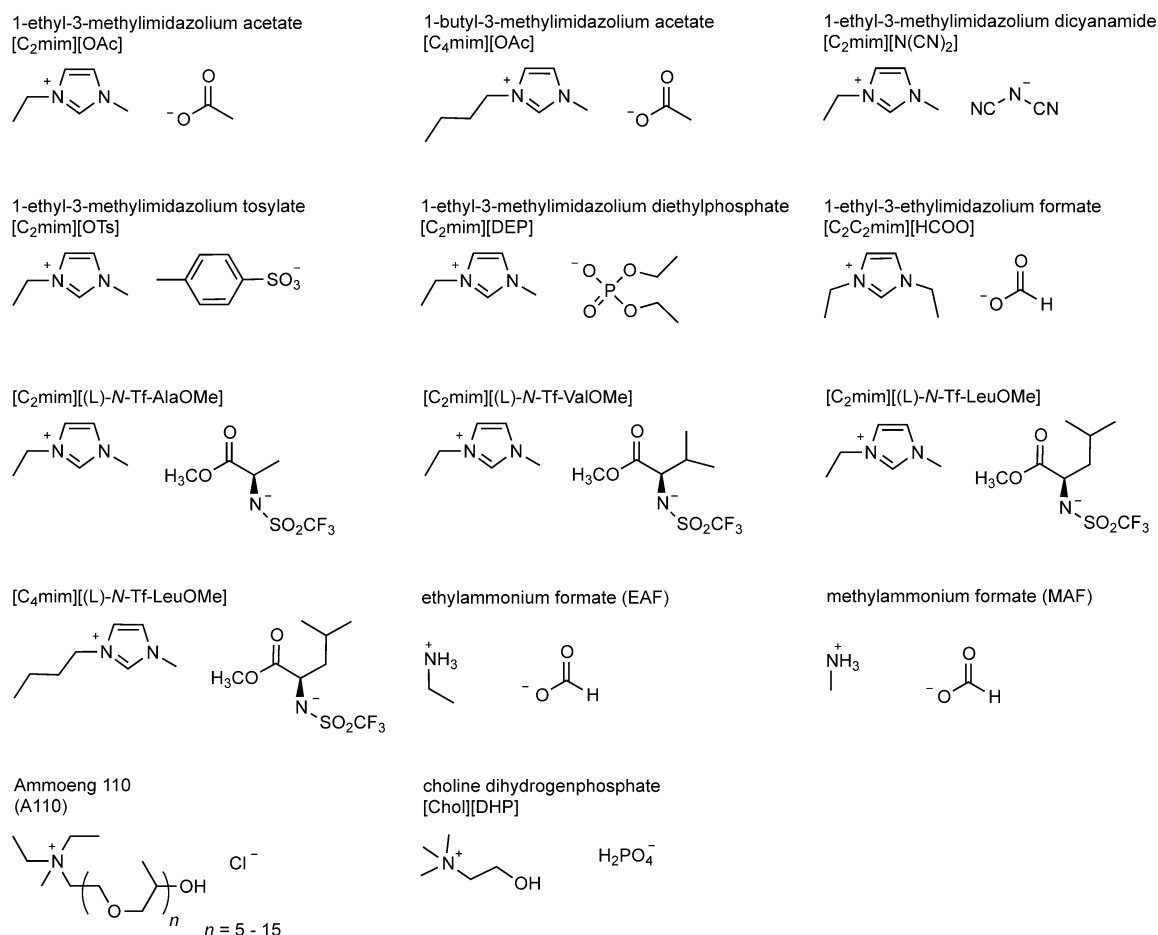
[d] Dr. A. Stark
Institute of Chemical Technology, University of Leipzig
Linnéstrasse 3-4, 04103 Leipzig (Germany)

[e] Dr. E. Leipold, Prof. Dr. S. H. Heinemann
Department of Biophysics, Center for Molecular Biomedicine
Friedrich Schiller University Jena and Jena University Hospital
Hans-Knöll-Strasse 2, 07745 Jena (Germany)

[f] Dr. C. Kandt
Computational Structural Biology, Life Science Informatics B-IT
LIMES Center, Mulliken Center for Theoretical Chemistry, University of Bonn
Dahmannstrasse 2, 53113 Bonn (Germany)

[†] These authors contributed equally to this work.

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201402356>.



Scheme 1. Structures of the ionic liquids used to synthesize conotoxin μ -SIIIA by the oxidative self-folding strategy, together with abbreviations.

have been successfully applied to chemical and enzymatic peptide syntheses, as well as fragment ligation.^[4,19] In this part of the study, the focus was placed on the use of 1) imidazolium-based RTILs differing in their anions (Scheme 1), and 2) conotoxin μ -SIIIA as a model peptide, although others such as α -GI, μ -PIIIA, and δ -EVIA were also included in the primary experiment. The plan to use imidazolium-based ILs with different anions was based on the idea that acetate as the anion supports the formation of the native conotoxin fold.^[14–16] It has been postulated that, thanks to its small size and basic character, it can participate in the disruption of hydrogen bonds between polymer chains. This, in turn, led us to hypothesize a similar mechanism with respect to the peptides investigated. However, whether or not this could be verified by the application of several other anions in combination with the 1-ethyl-3-methylimidazolium cation was one specific aim of this study.

In addition, the set of ILs was extended to further representatives (Scheme 1) that have recently attracted attention in peptide and protein research: namely Ammoeng 110, choline dihydrogenphosphate, and the protic ILs (PILs) methylammonium formate (MAF) and ethylammonium formate (EAF).^[20–24] The protic IL triethylammonium acetate (TEAA), for example, was applied in a study of the denaturation of the proteolytic enzyme α -chymotrypsin (α -CT) in the presence of urea. Attri

and co-workers described the attenuation of this process mediated by the non-ionic chaotrope urea through TEAA.^[22] In that work, the authors sought to address the counteracting effects of biocompatible ILs on urea-induced denaturation of biomolecules in comparison with their stabilizing effect on native protein folds and their function as refolding agents for α -CT. The PILs MAF, EAF, and butylammonium formate (BAF), on the other hand, were used in a study employing ultrafast transient absorption experiments to determine the local polarities of these PILs and their mixtures with water, with 12'-apo- β -carotene-12'-carboxylic acid (12'CA) as the molecular probe.^[23] Earlier experiments had suggested that the polarities of these ILs are comparable to, or even higher than, that of methanol or, in the case of MAF, even approaching that of water.^[23–25] In addition, PILs were reported to be more polar than most imidazolium-based ILs,^[26,27] thus making them interesting candidates for our study with respect to a direct comparison with the—so far—best working imidazolium-based IL, [C₂mim][OAc], for oxidative folding of conopeptides.^[14–16] However, it was speculated that further factors other than polarity, such as dipolarity and hydrogen-bonding effects, could influence their features and behavior in contact with a solute.

The study with the described selection of ILs, with μ -SIIIA as model, was further extended by the following aspects: firstly,

for the most promising ILs, reaction conditions (temperature, water content, reaction time) were varied, and, secondly, electrophysiological experiments were performed with μ -SIIIA products obtained from a large-scale synthesis in $[C_2mim][OAc]$. Interestingly, the large-scale synthesis had a significant effect on bioactivity, which was diminished relative to that of the product obtained by conventional synthesis in buffer systems; this is discussed with regard to solute–solvent interactions. In contrast, an enzymatic assay employing trypsin for proteolytic degradation of differentially oxidized μ -SIIIA versions revealed that the origin of the oxidized product, either from an aqueous medium or from an IL, leaves both the accessibility for enzymatic hydrolysis and the enzyme's catalytic activity virtually unaffected.

Results and Discussion

Effect of the anion of the ionic liquid

Oxidation of selected naturally occurring conopeptides^[28–33] (Figure S1 in the Supporting Information, Table 1) differing in

Table 1. Peptides selected for this study and corresponding yields for oxidation performed in different ionic liquids.

Reaction medium	α -GI	Oxidation yield [%]		
		μ -SIIIA	μ -PIIIA	δ -EVIA
buffer system ^[a]	$\approx 80^{[b]}$	≈ 50	16–18 ^[b]	$< 10^{[b]}$
$[C_2mim][OAc]$	≈ 50	≤ 70	≤ 60	25
$[C_4mim][OAc]$	< 5	n.d. ^[c]	38	< 5
$[C_2mim][OTs]$	n.p. ^[d]	n.p.	n.p.	n.p.
$[C_2mim][DEP]$	n.p.	n.p.	n.p.	n.p.
$[C_2mim][N(CN)_2]$	n.p.	n.p.	n.p.	n.p.

[a] Buffer solution contained 0.1 mM Tris-HCl (pH 8.7), 1 mM EDTA, 2 mM GSH, 1 mM GSSG, and 30% isopropanol, except in the case of δ -EVIA, in which 40% isopropanol was used. [b] Final peptide concentrations for redox-buffer systems were 0.05 mM (α -GI, μ -SIIIA, μ -PIIIA) and 0.02 mM (δ -EVIA). [c] The IL elutes with the same retention time as the oxidized product. The yield could not be determined with sufficient certainty. n.d.: not determined. [d] n.p.: no product was formed.

chain length (13–32 amino acids), cysteine content (four to six residues), cysteine pattern, net charge, and physicochemical properties were subjected to oxidative folding in the imidazolium-based RTILs $[C_2mim][OAc]$, $[C_4mim][OAc]$, $[C_2mim][OTs]$, $[C_2mim][DEP]$, and $[C_2mim][N(CN)_2]$, and comparison was made with the buffer systems described earlier.^[29–35] All compounds investigated were soluble in these ILs independent of their hydrophobicity. As a result, oxidation was carried out in rather high concentrations (7–12 mM; 1 mg of each precursor peptide was used) relative to the previously used aqueous buffer systems (0.02–0.05 mM),^[29–35] and in the absence of organic solvents, as well as at room temperature (vs. 4 °C commonly used for oxidation). Moreover, it was intended that the folding should proceed without any auxiliary redox agent [e.g., reduced and oxidized glutathione (GSH and GSSG, respectively)], similarly to our initial studies.^[14,15] With the exception of α -GI—

the simplest peptide in this study—the highest yield and quality of the peptide were obtained with $[C_2mim][OAc]$. This confirmed our earlier findings, yet we were surprised by the fact that no reaction occurred in the cases of the ILs with the anions $[OTs]^-$, $[DEP]^-$, and $[N(CN)_2]^-$ with either peptide. In contrast to other peptides, α -GI showed behavior in $[C_2mim][OAc]$ different from that found for the oxidation in aqueous buffer. This peptide contains only four cysteine residues, so only two disulfide bridges need to be formed upon oxidation. This particular folding might indeed proceed more easily than that of the other conotoxins, with three disulfide bridges and lower water solubility, and production in the aqueous medium is thus simple and unproblematic.

The set of peptides described here, together with the earlier findings for other peptides,^[15,16] led us to conclude that RTILs possessing anions of relatively high hydrogen bond acceptor strength and low degree of steric hindrance, such as acetate,^[36–38] support the folding of positively charged peptides (Table 1, Figure S1). In this type of IL, the oxidation is selective, resulting in higher yields than in conventionally used systems. However, comparison of these results with the Kamlet–Taft hydrogen bond acceptor strength (expressed as β values), as determined by Lungwitz and Spange, shows that the yields of the oxidative folding are independent of the β values, which are 0.85, 0.64, and 0.99 for $[C_4mim][OAc]$, $[N(CN)_2]$, and $[DEP]$, respectively.^[36] Changing the cation from $[C_2mim]^+$ to $[C_4mim]^+$ had no significant effect on solubility but decreased the overall yield.

Further imidazolium-based ILs (Scheme 1) with anions derived from amino acid derivatives (H-Xaa-OMe, with Xaa = Ala, Leu, Val) were used in combination with μ -SIIIA, the conotoxin obtained in the highest quantity and quality in the initial experiment, in order to evaluate their potential with regard to peptide folding. Surprisingly, the oxidized version of μ -SIIIA was not formed in any of those ILs (Figure S2). A closer look at the elution profiles also revealed different behavior of the peptide in these media, with peak splitting and broadening being observed. However, all of the major peaks displayed the same molar mass—that is, the molar mass of the reduced peptide—in MS analysis. It seems that in ILs with large anions (larger than acetate) peptide oxidation is hampered; these ILs even seem to stabilize the reduced form (i.e., the linear peptide). The peptide is thus probably not flexible enough to undergo folding, due to strong contact and interaction with the IL anions.

In an attempt to estimate the potential of the highly selective $[C_2mim][OAc]$ versus other imidazolium-based ILs for the synthesis of conotoxins more precisely, the reaction was additionally performed in binary ionic liquid mixtures also containing $[C_2mim][DEP]$. Here, the effect of a “non-folding” IL on the efficiency of peptide conversion in a “folding” IL—that is, $[C_2mim][OAc]$ —was monitored. The ILs ($[C_2mim][OAc]$ and $[C_2mim][DEP]$) were used in the following ratios: 1:3, 1:1, and 3:1 (v/v).

The results are shown in Figure 1: the higher the concentration of the $[DEP]^-$ anion in the reaction mixture, the lower was the yield. This behavior contrasts with results obtained in

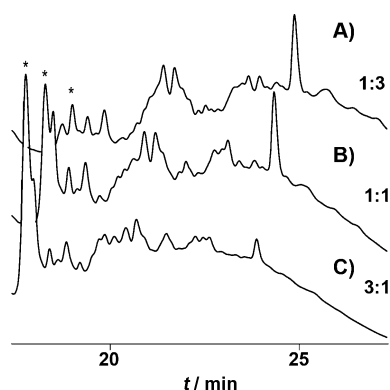


Figure 1. Monitoring of oxidative folding of μ -SIIIA in mixtures of $[\text{C}_2\text{mim}][\text{OAc}]$ and $[\text{C}_2\text{mim}][\text{DEP}]$ in different ratios: A) 1:3, B) 1:1, and C) 3:1 (v/v). HPLC conditions were as follows: gradient 0–40% eluent B, with eluent A being 0.1% TFA in water and eluent B being 0.1% TFA in acetonitrile; detection was at 220 nm (*: oxidized μ -SIIIA).

aqueous buffer solutions, in which the deprotonation of the cysteine thiol group is favored under more basic conditions ($\text{pH} > 8.0$), thus improving disulfide bridge formation.^[30,31,34]

Studies with ILs not yet applied to oxidative folding of Cys-rich peptides

In continuation of the initial experiments shown above, we focused our attention on studies employing ILs recently applied to proteins: that is, Ammoeng 110,^[20] choline dihydrogenphosphate (2 M in water),^[21] and selected protic ILs (MAF and EAF)^[22–25] (Figure 2). Moreover, $[\text{C}_2\text{C}_2\text{im}][\text{HCOO}]$ was also included, in order to cover two aspects: firstly the marginal variation of the cation arising from addition of one methylene group at position 3 of the imidazolium ring relative to $[\text{C}_2\text{mim}]^+$, and secondly the use of the formate anion as occurring in the protic ILs used in this study and displaying a basicity and structural make-up similar to those of acetate. In contrast with earlier findings with proteins,^[20,21,39–41] neither use of Ammoeng 110 nor that of choline dihydrogenphosphate led to the desired product. In the case of the latter medium, the fact that it needs to be dissolved in water to provide a liquid medium for synthetic application obviously was decisive for the negative result obtained. Indeed, this is in agreement with our findings that a water content of $> 3\%$ (v/v, see below) in the IL negatively influences the oxidation efficiency. Moreover, this IL is slightly acidic, due to the acidity of the H_2PO_4^- anion.^[21b] Ammoeng 110, on the other hand, has often been used previously in aqueous two-phase systems in connection with protein research.^[41] The majority of studies with proteins and ILs have also applied mixtures with rather high concentrations of water because many proteins lose their activity with decreasing water content.^[19,41]

Different results were obtained from the reactions performed in the protic ILs MAF (6% water, v/v) and EAF (3%, v/v). For both ILs the oxidized version of μ -SIIIA (0.5 mg, 11.3 mM) was identified, as demonstrated in Figure 2A. In contrast with the reaction performed in $[\text{C}_2\text{mim}][\text{OAc}]$, which re-

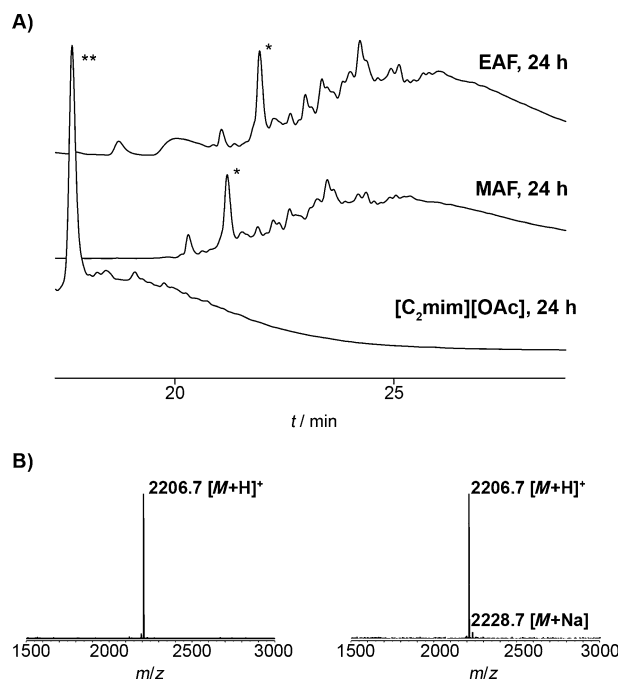


Figure 2. A) The protic ILs MAF and EAF were used for oxidation of linear μ -SIIIA over different time periods at 20 °C. The results shown (crude reaction mixtures) are those obtained after 24 h and compared with the same reaction in $[\text{C}_2\text{mim}][\text{OAc}]$ (bottom trace). RP-HPLC monitoring was performed with the gradient 0–40% eluent B, with eluent A being 0.1% TFA in water and eluent B being 0.1% TFA in acetonitrile; detection was at 220 nm. B) Mass spectra of oxidized μ -SIIIA (*) obtained from the reactions in MAF (left) and in EAF (right). The mass spectrum of the peak of oxidized μ -SIIIA (**) from $[\text{C}_2\text{mim}][\text{OAc}]$ is identical to the examples shown (see also ref. [14]).

vealed a major peak followed by a small broad and unresolved peak (see also Figure 3, below),^[14] yields could not be determined with any high degree of certainty when the protic ILs were used (see next paragraph). The quality of the product here was much lower than that of the product obtained in the imidazolium-based IL (retention times differ in these ILs). Surprisingly, the fractions isolated after elution at acetonitrile concentrations from 20 to 24% showed identical MS spectra (see, e.g., Figure 2B) indicating completely oxidized μ -SIIIA. The difference in the profiles obtained for MAF and EAF is likely to be due to the different water contents.

Finally, the use of $[\text{C}_2\text{C}_2\text{im}][\text{HCOO}]$ was also not successful in product formation (data not shown), similarly to what had been found for the previously mentioned amino-acid-based ILs. Therefore, the negative effect of the more highly alkylated cation on peptide oxidation, as found for $[\text{C}_4\text{mim}]^+$ in comparison with $[\text{C}_2\text{mim}]^+$, was confirmed.

The question of the cause of the differences between the PILs and the imidazolium-based IL remains. It is likely that several factors need to be taken into consideration, including the size of the cation (alkyl ammonium vs. imidazolium), protonation of peptide side chains triggered by the PIL cation, and the basicity of the anion (formate vs. acetate). Also, an opposite effect of these factors leading to partial compensation of each other cannot be excluded. Here in particular, a probable positive effect of the smaller cation in the PILs might be negatively

influenced by the lower basicity of the formate anion relative to acetate.

Variation of reaction conditions

To test the impact of water content (3–20%), reaction temperature (20–80 °C), reaction time (0.5–24 h), and scale (15–20-fold relative to preliminary experiments) on the yield and quality of μ -SIIIA, [C₂mim][OAc], MAF, and EAF were screened. Table 2

Table 2. Yields for the formation of oxidized μ -SIIIA (11.3 mM) obtained after 24 h reaction time. The reaction was generally monitored over a time period from 0.5 to 24 h.

Medium/conditions	[C ₂ mim][OAc]	Oxidation yield [%]	
		MAF ^[a]	EAF ^[a]
3 % water ^[b]	70	n.d. ^[c]	n.d.
5 % water	60	n.d.	56 (4 % water)
7 % water	40	40	40
10 % water	30	57 (12 % water)	42
20 % water	n.d.	36	30
20 °C	70	43	64
40 °C	70	48	50
60 °C	90/60 (6 h)	(69, 3 h) ^[d]	(58, 1 h) ^[d]
80 °C	90 (6 h)	(29, 1 h) ^[d]	(38, 1 h) ^[d]

[a] The peaks (elution at 20–24% of acetonitrile with 0.1 % TFA) containing completely oxidized μ -SIIIA were used for the determination of these reaction yields. [b] Reactions with different concentrations of water (v/v) in the IL were all carried out at 20 °C. [c] n.d.: not determined. [d] Peptide was degraded in this medium after the reaction time given.

summarizes the results of these tests at a peptide concentration of 11.3 mM (1 mg of peptide). The oxidized product was formed in the majority of these experiments; however, as can be seen in Figure 3, an increase in the water content led to a reduction in product yield. Surprisingly, this was found both for the imidazolium-based IL and for the protic ILs. When the concentration of water was low (\approx 3%) the oxidized product

was formed already after 6 h, whereas at higher concentrations (>5%) the reaction rate decreased and the product was not efficiently formed after 6 h reaction time. In [C₂mim][OAc], the finding of an improved stability of the peptide at higher temperatures than on treatment in aqueous solutions correlates well with observations for proteins, which were shown to refold to their native conformations without denaturation after heating to extreme temperatures (up to 110 °C).^[10,42] Full conversion of the linear peptide in [C₂mim][OAc] at 80 °C could be observed after just 3 h of reaction, unlike at lower temperatures (Figures 3B and S3). In contrast, no positive effect of the addition of water to the corresponding IL was observed for peptide oxidation (Figure 3A).

Interestingly, this is different from what has been reported for proteins,^[10] because the activities of these biomolecules depend to a certain individual degree on water in the operating system. This was explained by the fact that for an enzyme the so-called “essential” water (in contrast with the “bulk” free water), which is bound to the molecule in the form of several single layers on the surface, is critical for its activity. This water can be removed by hydrophilic ILs, whereas hydrophobic ILs show only a marginal tendency to work in this way. In more detail, in aqueous solution a protein’s functional groups are surrounded by a hydration shell, composed of water molecules attached to the charged groups on the protein surface; that is, water is essential for protein stability and activity. The hydrophobicity and polarity of the protein’s environment (e.g., solvent) thus influence the stability and activity by changing the protein hydration level.^[44] A pure IL is characterized by a hydrogen-bond network that is disrupted and modified in the presence of water in the reaction mixture.^[45] This indicates that the solution’s composition and “structure” changes in aqueous IL solutions.^[46] Upon interaction of IL ions with the protein surface, the protein can be stabilized in solution by a reduction in the protein–water interfacial tension through the “salting-out” effect, but with an increasing water concentration the hydration forces start to dominate, and the ions start to confer a “salting-in” effect.

In comparison with earlier reports,^[14,15] we noticed that an increase in the batch size (\geq 15 mg of peptide, 11.3 mM) requires an additional supply of air oxygen and careful stirring for a successful reaction (i.e., mass transfer limitations occur otherwise). Further studies directed towards up-scaling of the oxidation of reduced μ -SIIIA (15–20 mg) in [C₂mim][OAc] were primarily focused on the impact on reaction rate, efficiency, and yield.

Indeed, we observed a dramatic influence of these parameters on the oxidative folding process and the product obtained. The results acquired after 24 h reaction time are shown in Figure 4 and revealed an influence of the ionic liquid ions on the peptide’s elution behavior. Whereas the buffer-oxidized version of μ -SIIIA (1) eluted at 17.3% acetonitrile (0.1 % TFA, v/v), the buffer-oxidized IL-incubated peptide eluted at 22–24% (peaks 2a and 2b).^[14] However, the chromatogram of μ -SIIIA obtained from the scaled-up approach contained peaks at 16.5 min (i.e., 16.5% of 0.1 % TFA in acetonitrile; fraction 3), 17.6 min (fraction 4), 20.8 min (fraction 5), 22.9 min (fraction 6),

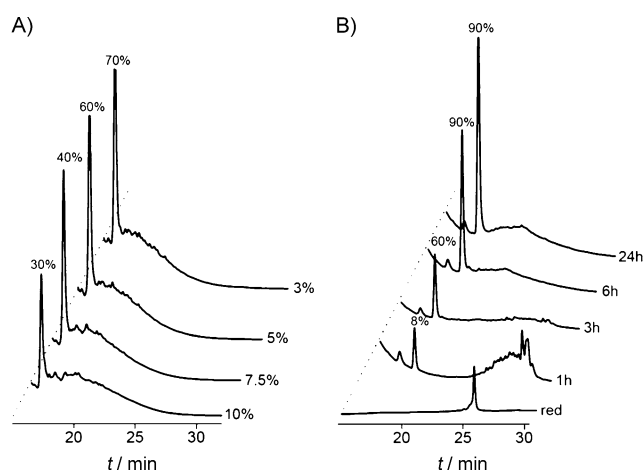


Figure 3. Elution profiles for oxidative folding of μ -SIIIA (11.3 mM) in [C₂mim][OAc], depending on: A) water content after 24 h at room temperature, and B) reaction time at 80 °C and with 3 % water content (HPLC conditions as in Figure 2).

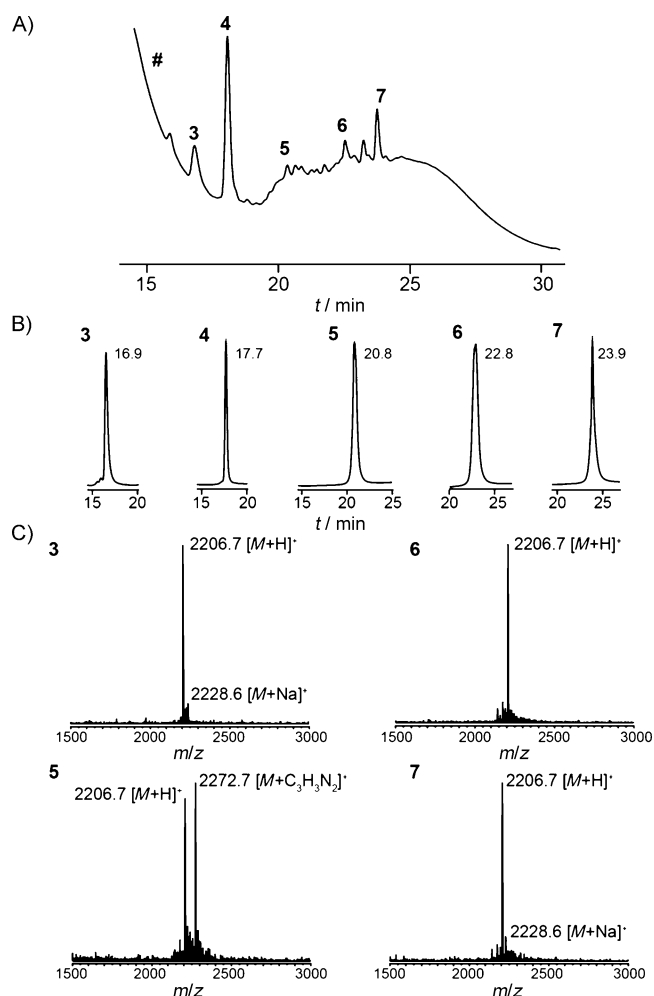


Figure 4. A large amount of reduced μ -SIIIA (15–20 mg) was oxidized in [C₂mim][OAc] (#: peak of IL) at 20 °C over 24 h. A) Beside the main peak (fraction 4 several other peaks were observed, and the corresponding products were isolated from the crude reaction mixture. B) Products corresponding to individual peaks (fractions 3 and 5–7) were reinjected after separation of the crude mixture (retention times are given), and C) subjected to MS analysis with a MALDI-TOF mass spectrometer. An MS spectrum corresponding to fraction 4 was published earlier.^[14] HPLC conditions were as described in Figure 2.

and 23.6 min (fraction 7), all showing the same molar mass (2208.7 g mol⁻¹, [M+H]⁺ of oxidized peptide) in MS analysis (Figure 4, inserts). It seemed that ionic liquid cations and anions adhere to the peptide, leading to varying forms with respect to ion composition and amount. It was supposed that these ions are not covalently or strongly attached because MALDI-TOF MS analysis did not reveal any of these compositions (Figure 4B). Whether these variants represent different μ -SIIIA isomers or just one single isomer with different ion compositions was thus investigated in more detail by MS/MS analysis (see below).

The complete transformation of the reduced precursor into the oxidized peptide was generally confirmed by mass spectrometry and Ellman's test (Table S1). In addition, Raman spectroscopy (Figure S4) was used to verify the completion of the folding process. In contrast with time-consuming and yet un-

suitable NMR analysis of the rather complex peptide, Raman spectroscopy is a fast and convenient method to derive the required information in addition to MS and has previously been successfully applied in the same respect.^[14, 15, 43]

Electrophysiological studies

In continuation of our previous studies on the subject,^[14] we measured the activities of different μ -SIIIA fractions, produced by various procedures, in blocking voltage-gated sodium channels Na_v1.4 (Figure 5). In addition to the peptide oxidized in

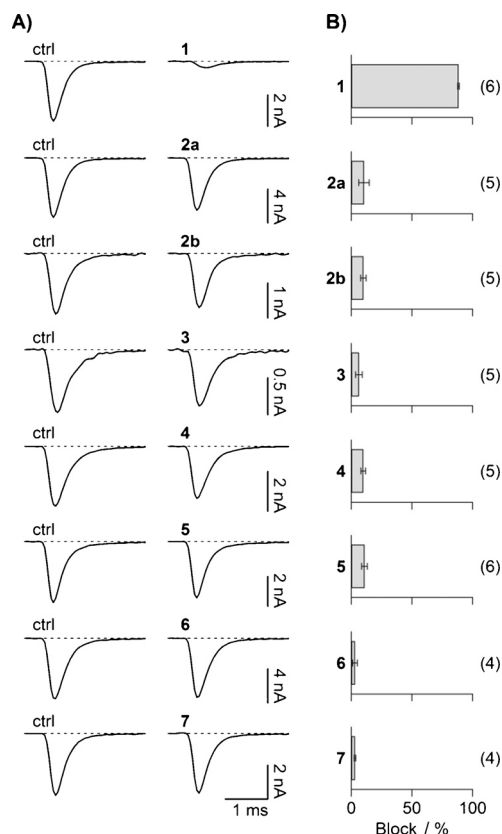


Figure 5. Blocking of sodium channel Na_v1.4 by different versions of oxidized μ -SIIIA. A) Na_v1.4 was heterologously expressed in HEK293 cells, and current responses were recorded at test depolarizations to 0 mV before (left, ctrl) and after (right) application of the indicated peptide fraction for a period of 500 s. Fraction 1: buffer-oxidized peptide (10 μ M); 2a and 2b: buffer-oxidized and treated with [C₂mim][OAc] (90 μ M); 3–7: various fractions purified after oxidation in [C₂mim][OAc] (all 90 μ M) obtained from the experiment shown in Figure 4. B) Average activities of the indicated peptide fractions in blocking Na_v1.4 channels. Data are presented as means \pm SEM, with the number of independent experiments indicated in parentheses.

a buffer system (Figure 5, fraction 1),^[14] we tested the activity of buffer-folded μ -SIIIA after additional incubation (24 h) in [C₂mim][OAc] (Figure 5, fractions 2a and 2b). As mentioned above, IL treatment changes the HPLC retention time of the peptide. Moreover, we found an influence of this treatment on the peptides' ability to inhibit Na_v1.4 channels. In contrast with the blocking induced by 10 μ M buffer-oxidized μ -SIIIA (fraction 1, ca. 84%), the IL-treated peptide of fraction 2, at a con-

centration of 90 μM , showed only 11% (**2a**) and 10% (**2b**). A similar reduction in activity was found for the fractions purified from a large-scale oxidation in $[\text{C}_2\text{mim}][\text{OAc}]$: fractions **3** (6%), **4** (10%), **5** (11%), and almost no detectable activity in **6** (3%) and **7** (3%; 90 μM each, Figure 5).

In this respect, there was no difference between fractions obtained from analytical or semipreparative RP-HPLC separation of the buffer-oxidized and $[\text{C}_2\text{mim}][\text{OAc}]$ -treated peptide and the peptide oxidized in $[\text{C}_2\text{mim}][\text{OAc}]$.

The results found here contrast with what we had observed for the small-scale-produced peptide and also with earlier studies with an enzyme inhibitor, also synthesized in $[\text{C}_2\text{mim}][\text{OAc}]$.^[14,16] It is generally recognized that the up-scaling of a process can be accompanied by difficulties. The fact that products with differing retention times but with the same molar mass and similar activity were observed reveals a different behavior and mechanism with regard to solute–IL interactions for the up-scaling and purification process relative to that described previously. Here, it is obvious that HPLC purification was not sufficient to remove the noncovalently bound IL ions from the peptide molecule. LC-ESI MS/MS analysis was used to clarify the disulfide connectivity by application of a protocol of partial reduction combined with carbamidomethylation (Supporting Information). The disulfide bridges of the buffer-oxidized peptide were first analyzed by this procedure and predominantly showed the suggested native folding: Cys3–Cys13, Cys4–Cys19, Cys8–Cys20 (Table S2, Figures S8 and S9). Additionally, minor contamination with another isomer (i.e., Cys3–Cys8, Cys4–Cys13, Cys19–Cys20) was observed. Subsequently, the same method was applied to investigate the connectivity of the buffer-oxidized $\mu\text{-SIIIA}$ incubated with $[\text{C}_2\text{mim}][\text{OAc}]$; this gave the same results, clearly with no changes due to the IL (Figures S10 and S11). Interestingly, the isomer corresponding to the native fold was identified in fraction **6** of the IL-oxidized $\mu\text{-SIIIA}$ (Table S2, Figures S12–S14). In addition, fraction **5** also predominantly showed the same MS/MS fragmentation: that is, the native fold but with hints of a minor portion of non-native disulfide bridging (Table S2). In contrast, fractions **4** and **7** represent a disulfide connectivity different from that in fractions **5** and **6**. However, there was no single, dominant connectivity found for these versions. Several different linkages hinted at, for example, Cys4–Cys8, Cys3–Cys19, and Cys8–Cys13 (data not shown). Nevertheless, the shift of the retention times of fractions **5** and **6** relative to the buffer-oxidized $\mu\text{-SIIIA}$ is due to the adhesion of ions to the cyclic peptide. As a consequence, the loss of bioactivity could be attributable to these ions if the interactions of the most relevant amino acids of the peptide with the ion channel are disturbed. To verify this hypothesis, molecular dynamics simulations were performed (see below).

Impact of IL ions on proteolytic degradation of $\mu\text{-SIIIA}$

To examine the influence of the oxidation procedure on the quality and activity of $\mu\text{-SIIIA}$, we performed an enzymatic assay with the peptide as competing substrate. The intention was to compare different kinds of activity assays and, thus,

sensitivities of biological targets (enzyme vs. ion channel) to the IL ions adhering to the peptide. Proteolytic degradation by trypsin was examined in the presence of both oxidation products: the toxin obtained from aqueous buffer and the version from the $[\text{C}_2\text{mim}][\text{OAc}]$ approach. Trypsin activity was monitored with the aid of a chromogenic substrate ($N\alpha$ -benzoyl-DL-arginine-4-nitroanilide hydrochloride, BAPNA) by measuring the increase in absorption at 410 nm. $\mu\text{-SIIIA}$ contains three potential cleavage sites for trypsin (K11, R14, R18). In the presence of $\mu\text{-SIIIA}$, the peptide competes with the substrate BAPNA for the active site of trypsin, thereby lowering the absorption increase and the “apparent” activity of trypsin. The quantity of this effect was taken as a measure to assess the cleavage of $\mu\text{-SIIIA}$ by trypsin. For buffer-oxidized $\mu\text{-SIIIA}$, we derived a concentration-response curve ($\text{IC}_{50} = 6.46 \pm 1.60 \mu\text{M}$, Figure 6), and the maximum inhibition of BAPNA cleavage was

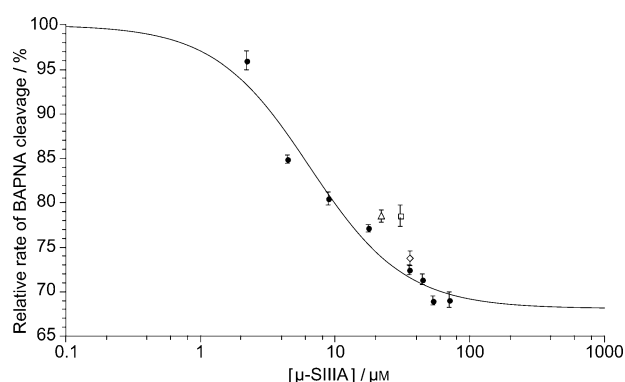


Figure 6. Proteolytic cleavage of substrate BAPNA in the presence of $\mu\text{-SIIIA}$ from different oxidation procedures. Buffer-oxidized $\mu\text{-SIIIA}$ (●); IL-oxidized $\mu\text{-SIIIA}$ **5** (□), **6** (◇), **7** (△).

determined to be 68% ($\pm 3\%$) of the value obtained in the absence of $\mu\text{-SIIIA}$. Selected fractions of the IL-oxidized $\mu\text{-SIIIA}$ were then tested at a distinct concentration (from the stock solutions used in electrophysiological experiments) and used in the same way as the buffer-oxidized peptide.

With regard to the effect on BAPNA conversion these peptide versions showed behavior similar to that of the buffer-oxidized $\mu\text{-SIIIA}$, for which a reduction of 72% at the tested concentration was detected. The results for fractions **5** ($\approx 78.5\%$), **6** ($\approx 73.8\%$), and **7** ($\approx 78.5\%$) differ slightly, but they were in an acceptable range for the assay setup used ($\pm 1.8\text{--}6.5\%$ deviation; Figure 6). Nevertheless, in relation to what has been observed in the electrophysiological experiment with a $\geq 70\text{--}80\%$ deviation of the results for the IL-oxidized peptide relative to the buffer-oxidized $\mu\text{-SIIIA}$, the impact of the medium used in the oxidation procedure on the peptide's activity is not as large for the enzyme as for the ion channel. This indicates a much higher sensitivity of the latter and the corresponding toxin–protein interaction than for the enzyme–substrate interaction, probably based on differences in binding area and number of specific amino acid residues involved in the recognition process on either side.

Hotspots of IL ions interacting with μ -SIIIA

Using atomistic molecular dynamics simulation, we sampled the μ -SIIIA- $[\text{C}_2\text{mim}][\text{OAc}]$ interaction on a 100 ns timescale, and for each residue we determined the average amount of ≤ 2.5 Å contacts to $[\text{C}_2\text{mim}]^+$ (Figure 7A and C, Figure S15A and C) and $[\text{OAc}]^-$ (Figure 7B and D, Figure S15B and D). For

peptides or proteins can, however, only be provided through greatly increased research activities in this field and the acquisition of corresponding experimental data. Therefore, a selection of ionic liquids has been used here to explore the impact of the IL medium on the formation of oxidatively folded peptides, in particular the conotoxin μ -SIIIA. It has been shown that the oxidized peptide was formed to a relevant extent in protic ILs,

whereas most imidazolium-based ILs (differing in their anions), as well as choline dihydrogenphosphate and Ammoeing 110, failed to yield the product. Nevertheless, in contrast with the results obtained for the only efficiently functioning 1,3-dialkylimidazolium IL— $[\text{C}_2\text{mim}][\text{OAc}]$ —the protic ILs proved to give highly complex product mixtures, thus making them unsuitable as adequate substitutes for established aqueous systems.

Another previously unobserved aspect was seen upon increasing the batch size. Instead of the expected transfer of the reaction outcome in terms of quality and quantity for the scaled-up approach in $[\text{C}_2\text{mim}][\text{OAc}]$, a rather unusual mixture was obtained and analyzed. Isolated materials with different physico-chemical properties with respect to the behavior in the chromatographic separation process, yet with identical molar masses, displayed massive reductions in biological activity relative to the toxin prepared in

the aqueous medium. This is a direct consequence of favorable strong interactions between the peptide's backbone and amino acid side chains and ionic liquid cations and anions. The experiment with the IL-incubated toxin, which was active before the contact with the IL, the observations in LC-ESI MS/MS analysis, and the molecular dynamics simulation strongly support this idea. Although solutes generally interact with the reaction medium—water in the case of biological materials, for example—the findings of this study revealed a clear impact of the IL on the biological recognition process. Therefore, the contradictory findings relating to proteins in ILs described in the literature might be the result of IL-dependent interactions with the protein of interest disturbing the process to be analyzed. Fundamental research thus remains to be carried out to clarify the effects observed and to allow full exploitation of the potential of ILs for future applications in biological sciences.

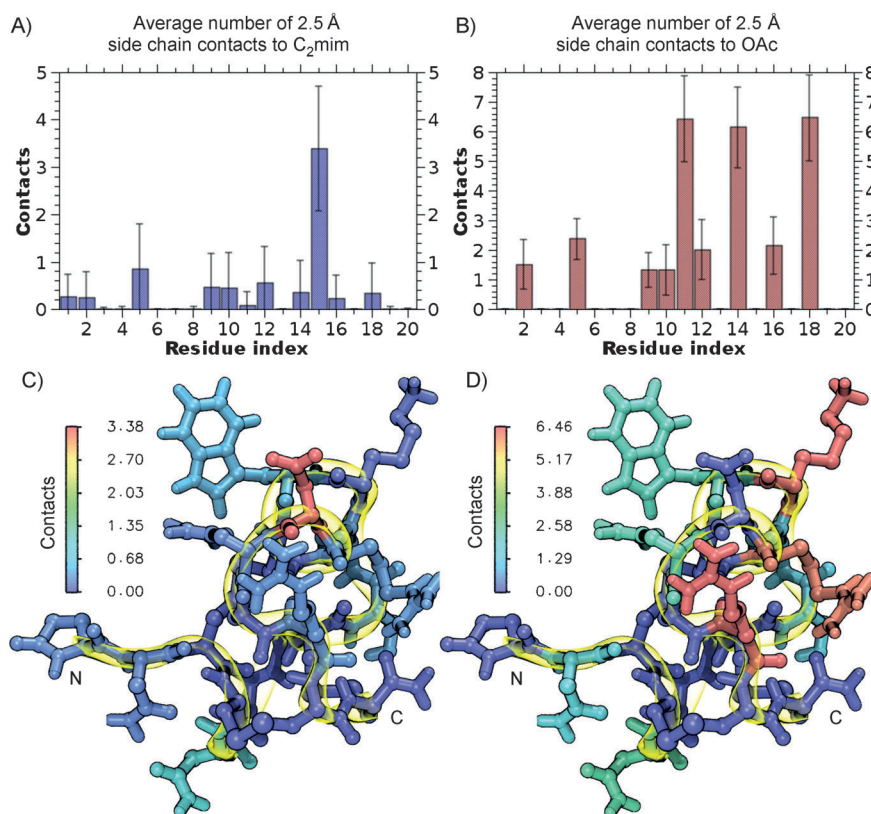


Figure 7. Hotspots of $[\text{C}_2\text{mim}][\text{OAc}]$ - μ -SIIIA interaction as determined by molecular dynamics simulations. For each residue the average amount of close contacts to A), C) $[\text{C}_2\text{mim}]^+$, and B), D) $[\text{OAc}]^-$ is shown quantitatively and in the context of a μ -SIIIA conformation closest to the simulation average structure. C), D) For clarity, the course of the peptide backbone has been highlighted, with N and C marking the termini.

both folded (Figure 7) and unfolded μ -SIIIA (Figure S15) we found that $[\text{C}_2\text{mim}]^+$ interactions predominantly take place at Asp15, whereas close contacts with the anion preferentially occur at Lys11, Arg18, and Arg14. Together with the MS/MS analysis of the disulfide connectivities of the different peptide versions—that is, identification of the same isomers in differentially oxidized μ -SIIIA—this is a clear indication of IL ions being responsible for activity loss, particularly if it is considered that these basic residues are primarily responsible for interaction with the ion channel pore.^[32]

Conclusion

Despite the recent progress achieved in the development and application of ILs for biocatalytic processes and for reactions involving proteins in general, research into employment of ILs in peptide chemistry is still in its infancy. Detailed understanding of the molecular interactions between ionic liquids and

Experimental Section

Chemicals: Fmoc-amino acids, coupling reagents (HBTU), and resin were purchased from Orpegen Peptide Chemicals (Heidelberg, Germany) or IRIS Biotech (Marktredwitz, Germany). Peptide synthesis reagents (*N*-methylmorpholine, piperidine, trifluoroacetic acid) and solvents (*N,N*-dimethylformamide, CH₂Cl₂) were of reagent grade, solvents for chromatography (methanol, acetonitrile) were of analytical grade and were obtained from VWR International (Dresden, Germany), and HPLC water was obtained from Fisher Scientific. [C₂mim][OAc] and [C₄mim][OAc] were donated by BASF SE and Sigma-Aldrich, respectively. [C₂mim][N(CN)₂], [C₂mim][DEP], [C₂mim][OTf], EAF, MAF, and Amoeng were obtained from IoLiTec, Solvent Innovation, or Merck. Other ionic liquids were synthesized as described below.

Synthesis of ionic liquids

1-Butyl-3-methylimidazolium bromide: 1-Bromobutane (1.2 equiv) was heated to 70 °C, and *N*-methylimidazole (1.0 equiv) was added dropwise. The solution was allowed to cool to room temperature, diluted with CH₂Cl₂, and poured dropwise with intensive mechanical stirring into cold ethyl acetate. The product separated immediately as a white precipitate. The solid was crushed to a powder and washed twice with ethyl acetate (200 mL) by decantation. The highly hygroscopic product was dried in vacuum to constant weight and stored under ethyl acetate. Colorless crystalline solid, 25.2 g (115 mmol, 75%), ¹H NMR (300 MHz, [D₆]DMSO): δ = 9.42 (s, 1H), 7.89 (s, 1H), 7.78 (s, 1H), 4.17 (t, ³J_{H,H} = 7.2 Hz, 2H), 3.86 (s, 3H), 1.66–1.73 (m, 2H), 1.12–1.26 (m, 2H), 0.82 ppm (t, ³J_{H,H} = 7.4 Hz, 3H); ¹³C NMR (75 MHz [D₆]DMSO): δ = 137.0, 122.7, 124.0, 48.8, 36.2, 31.8, 19.2, 13.8 ppm.

1-Ethyl-3-methylimidazolium bromide: 1-Bromoethane (1.0 equiv) was heated to 40 °C, and *N*-methylimidazole (1.0 equiv) was added dropwise. The reaction mixture was stirred at 40 °C for 1 h. Afterwards the temperature was increased to 80 °C, and the mixture was stirred for another hour. The solution was allowed to cool to room temperature and poured dropwise with intensive mechanical stirring into cold ethyl acetate. The product separated immediately as a white precipitate. The solid was isolated and dissolved in CH₂Cl₂. Activated charcoal and aluminium oxide were added, and the solution was filtered off. The filtrate was treated with ethyl acetate. After phase separation, the lower, yellow phase was isolated and dried under reduced pressure at 50 °C. After cooling to room temperature, the product crystallized slowly as a colorless solid. Colorless crystalline solid, 55.8 g (294 mmol, 73%). ¹H NMR (300 MHz, [D₆]DMSO): δ = 9.74 (s, 1H), 7.17–7.10 (m, 2H), 3.89 (q, ³J_{H,H} = 7.44 Hz, 2H), 3.58 (s, 3H), 1.06 ppm (t, ³J_{H,H} = 7.4 Hz, 3H); ¹³C NMR (75 MHz, [D₆]DMSO): δ = 136.2, 123.4, 121.9, 44.0, 36.2, 35.7, 15.2 ppm.

1-Ethyl-3-methylimidazolium hydroxide ([C₂mim][OH]): An aqueous solution was prepared from 1-ethyl-3-methylimidazolium bromide by use of anion exchange resin (Merck Ion exchanger III) and ultrapure water. Qualitative ion chromatography was used to check whether all halide impurities had been removed.

General procedure for *N*-trifluoromethanesulfonyl amino acid methyl esters: Triethylamine (2.0 equiv) was dissolved in abs. CH₂Cl₂ under argon. The L-amino acid methyl ester hydrochloride (1.0 equiv) was added to the solution, and the mixture was stirred for 5 min at room temperature. The solution was cooled to –78 °C (isopropanol/dry ice). Trifluoromethanesulfonic acid anhydride (1.0 equiv) dissolved in CH₂Cl₂ was added. The mixture was allowed to warm slowly to room temperature and stirred overnight. The so-

lution was treated with hydrochloric acid (9%), and the organic layer was separated off. The organic layer was washed with brine and concentrated to remove the solvents. The crude product was purified by Kugelrohr distillation or by sublimation.

(*S*)-*N*-(Trifluoromethanesulfonyl)alanine methyl ester [N-Tf-AlaOMe]: White crystalline solid, 1.85 g (7.84 mmol, 85%), Kugelrohr distillation at 120 °C and 0.01 mbar. M.p.: 63 °C; ¹H NMR (300 MHz, [D₆]DMSO): δ = 10.28–10.14 (s, 1H), 4.22 (m, 1H), 3.73 (s, 3H), 1.38 ppm (d, ³J_{H,H} = 7.3 Hz, 3H); ¹³C NMR (75 MHz [D₆]DMSO): δ = 172.0, 119.8 (q, ¹J_{C,F} = 321.3 Hz), 52.6, 53.0, 18.9 ppm; MS (ESI): *m/z* calcd for C₅H₉NO₄FS₃: 234.0 [*M*–1][–]; found: 233.8.

(*S*)-*N*-(Trifluoromethanesulfonyl)valine methyl ester [N-Tf-ValOMe]: Colorless crystalline solid, 2.63 g (9.99 mmol, 65%), Kugelrohr distillation at 120 °C and 0.01 mbar. ¹H NMR (300 MHz, [D₆]DMSO): δ = 10.06 (brs, 1H), 3.84 (m, 1H), 3.70 (s, 3H), 2.04–2.15 (m, 1H), 0.91 (d, ³J_{H,H} = 2.5 Hz, 3H), 0.89 ppm (d, ³J_{H,H} = 2.5 Hz, 3H); ¹³C NMR (75 MHz [D₆]DMSO): δ = 171.1, 121.9 (q, ¹J_{C,F} = 321.81 Hz), 62.9, 53.7, 30.7, 19.1, 18.1 ppm; MS (ESI): *m/z* calcd for C₇H₁₁NO₄FS₃: 262.23 [*M*–1][–]; found: 263.00.

(*S*)-*N*-(Trifluoromethanesulfonyl)leucine methyl ester [N-Tf-LeuOMe]: Colorless crystalline solid, 4.51 g (16.3 mmol, 72%), Kugelrohr distillation at 120 °C and 0.01 mbar. ¹H NMR (300 MHz, [D₆]DMSO): δ = 10.13 (brs, 1H), 4.02–3.97 (m, 1H), 3.68 (s, 3H), 1.50–1.62 (m, 3H), 0.89 ppm (t, ³J_{H,H} = 6.4 Hz, 6H); ¹³C NMR (75 MHz [D₆]DMSO): δ = 171.5, 119.3 (q, ¹J_{C,F} = 321.5 Hz), 55.1, 52.4, 40.3, 23.9, 22.52, 20.6 ppm; MS (ESI): *m/z* calcd for C₈H₁₃NO₄FS₃: 276.25 [*M*–1][–]; found: 275.9.

1-Butyl-3-methylimidazolium (*S*)-*N*-(trifluoromethanesulfonyl)alanine methyl ester [C₄mim][N-Tf-L-Ala-OMe]: The protonated anion precursor (1.0 equiv) was dissolved in CH₂Cl₂ under argon. Sodium hydride (60% in mineral oil, 1.7 equiv, effective dose 1.0 equiv) was added, and the solution was stirred until the evolution of gas had ceased. With continuous stirring, the desired ionic liquid cation (1.0 equiv) was added as a solution (1 M) of the bromide salt in CH₂Cl₂. The solution was stirred for 24 h at ambient temperature for grain growth by Oswald ripening. The solid was filtered off, and the remaining solution was washed with small portions of water until a test for halides with AgNO₃ was negative. The solvent was removed under reduced pressure, and the residue was dried to constant weight under vacuum. Pale yellow liquid, 1.40 g (3.75 mmol, 88%). M.p.: < 0 °C; ¹H NMR (300 MHz [D₆]DMSO): δ = cation: 9.42 (s, 1H), 7.87 (s, 1H), 7.79 (s, 1H), 4.20 (t, ³J_{H,H} = 7.3 Hz, 2H), 3.88 (s, 3H), 1.84–1.69 (m, 2H), 1.33–1.19 (m, 2H), 0.87 (t, ³J_{H,H} = 7.4 Hz, 2H), anion: 3.95–3.80 (m, 1H), 3.54 (s, 3H), 1.15 ppm (d, ³J_{H,H} = 7.3 Hz, 3H); ¹³C NMR (75 MHz [D₆]DMSO): δ = anion: 176.0, 122.6 (q, ¹J_{C,F} = 330 Hz), 54.3, 51.5, 22.1, cation: 137.3, 122.6, 124.2, 49.1, 36.2, 31.9, 19.3, 13.8 ppm.

General procedure for 1-ethyl-3-methylimidazolium (*S*)-*N*-trifluoromethanesulfonyl amino acid methyl ester ionic liquids: The (*S*)-*N*-trifluoromethanesulfonyl amino acid methyl ester (1.0 equiv) was dissolved in ultrapure water. A solution of 1-ethyl-3-methylimidazolium hydroxide (1.0 equiv) in ultrapure water was added dropwise over a period of 1.5 h. The resulting mixture was stirred at room temperature for 18 h. The water was removed under reduced pressure to afford the desired product as a highly viscous liquid.

1-Ethyl-3-methylimidazolium (*S*)-*N*-(trifluoromethanesulfonyl)alanine methyl ester [C₂mim][N-Tf-L-Ala-OMe]: Colorless liquid, 2.3 g (6.56 mmol, 77%). M.p.: < 0 °C; ¹H NMR (300 MHz [D₆]DMSO): δ = cation: 9.31 (s, 1H), 7.80 (s, 1H), 7.71 (s, 1H), 4.20 (q, ³J_{H,H} =

7.3 Hz, 2H), 3.85 (s, 3H), 1.40 (t, $^3J_{\text{H,H}} = 7.3$ Hz, 3H), anion: < 3.89–3.82 (m, 1H), 3.51 (s, 3H), 1.11 ppm (d, $^3J_{\text{H,H}} = 6.9$ Hz, 3H); MS (ESI): m/z calcd for $\text{C}_6\text{H}_{11}\text{N}_2^+$: 111.1 $[M-1]^+$; found: 111.0; m/z calcd for $\text{C}_5\text{H}_7\text{NO}_4\text{SF}_3^-$: 234.0 $[M-1]^-$; found: 233.9.

1-Ethyl-3-methylimidazolium (S)-N-(trifluoromethanesulfonyl)valine methyl ester [C₂mim][N-Tf-L-Val-OMe]: Colorless liquid, 1.2 g (3.33 mmol, 88%). M.p.: < 0 °C; ^1H NMR (300 MHz [D₆]DMSO): δ = cation: 9.22 (s, 1H), 7.78 (s, 1H), 7.70 (s, 1H), 4.19 (q, $^3J_{\text{H,H}} = 7.3$ Hz, 2H), 3.84 (s, 3H), 1.40 (t, $^3J_{\text{H,H}} = 7.3$ Hz, 3H), anion: 3.50 (s, 3H), 3.46 (d, $^3J_{\text{H,H}} = 6.6$ Hz, 1H), 1.72 (m, 1H), 0.78 ppm (t, $^3J_{\text{H,H}} = 6.7$ Hz, 6H); MS (ESI): m/z calcd. for $\text{C}_6\text{H}_{11}\text{N}_2^+$: 111.1 $[M-1]^+$; found: 111.0; m/z calcd for $\text{C}_7\text{H}_{11}\text{NO}_4\text{SF}_3^-$: 262.2 $[M-1]^-$; found: 262.9.

1-Ethyl-3-methylimidazolium (S)-N-(trifluoromethanesulfonyl)leucine methyl ester [C₂mim][N-Tf-L-Leu-OMe]: Colorless liquid, 1.2 g (3.11 mmol, 76%). M.p.: < 0 °C; ^1H NMR (300 MHz [D₆]DMSO): δ = cation: 9.36 (s, 1H), 7.78 (s, 1H), 7.69 (s, 1H), 4.20 (q, $^3J_{\text{H,H}} = 7.3$ Hz, 2H), 3.85 (s, 3H), 1.40 (t, $^3J_{\text{H,H}} = 7.3$ Hz, 3H), anion: 3.77 (t, $^3J_{\text{H,H}} = 7.1$ Hz, 1H), 3.49 (s, 3H), 1.61 (m, 1H), 1.29 (t, $^3J_{\text{H,H}} = 7.0$ Hz, 2H), 0.81 ppm (t, $^3J_{\text{H,H}} = 7.0$ Hz, 6H); MS (ESI): m/z calcd. for $\text{C}_6\text{H}_{11}\text{N}_2^+$: 111.1 $[M-1]^+$; found: 111.0; m/z calcd for $\text{C}_8\text{H}_{13}\text{NO}_4\text{SF}_3^-$: 276.1 $[M-1]^-$; found: 275.9.

Synthesis and purification of peptides: Peptides were synthesized automatically by a standard Fmoc [N-(9-fluorenyl)methoxycarbonyl] protocol as described earlier^[14,32] with an EPS 221 peptide synthesizer (Intavis Bioanalytical Instruments AG, Cologne, Germany). The linear peptides (conotoxin sequences are represented in Figure S1) were assembled by use of TentaGel R RAM resin with a loading capacity of 0.18 or 0.24 mmol g⁻¹. Cysteine side chains were protected with trityl groups. In general, coupling reactions were performed with Fmoc-amino acids (5 equiv) activated with HBTU (5 equiv) in the presence of *N*-methylmorpholine/DMF (1:1) for 5–15 min (double couplings). Fmoc removal was effected by treating the resin twice with piperidine in DMF (20%). All deprotection and coupling steps were followed by intensive washings with DMF and CH_2Cl_2 , alternately. Peptide cleavage and deprotection was accomplished with reagent K (trifluoroacetic acid (TFA)/water/phenol/thioanisole/ethanedithiol 82.5:5:5:2.5) for 4 h at room temperature. The crude peptides were precipitated with diethyl ether, centrifuged, and washed several times with diethyl ether. The yields of the crude peptides varied between 70–90% for the linear sequences. Peptides were purified by semipreparative reversed-phase C18 HPLC with the gradient elution system 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile/water (90:10) (eluent B).

Oxidation of peptides: Oxidative folding of peptides α -GI, μ -SIIIA, μ -PIIIA, and δ -EVIA in buffer with redox agents was performed by the procedure described earlier.^[14,32] In ILs (preliminary experiment), the linear peptide (1 mg) was dissolved in the corresponding IL (40–120 μL) and left at room temperature with gentle stirring. For experiments on the impact of water content in ILs, gradient grade water (VWR, Dresden, Germany) was added to the reaction mixtures prior to the start of the reaction in final concentrations of 3, 5, 7.5, 10, and 20%. For experiments on the impact of elevated temperatures, the corresponding reaction mixture was left in a thermomixer at 40, 60, and 80 °C. After an appropriate time (0.5–24 h), the mixture was diluted 16-fold with acetonitrile/water (1:4) and purified by RP-HPLC by the methods described in Table S1.

Characterization of ionic liquids: The water content of each ionic liquid was determined by a Karl Fischer titration (Metrohm 870 KF Titrino plus).^[47]

Molecular dynamics simulations: MD simulations were performed with GROMACS version 4.0.3^[48,49] and the GROMOS96 force field with the 54a7 parameter set^[50] by using Vienna PTM to model conotoxin's post-translationally modified N and C termini.^[51,52] In all simulations standard protonation states were assumed for titratable residues and all bond lengths to hydrogen were constrained by LINCS^[53] so that an integration time step of 2 fs could be chosen. Systems were simulated at 300 K, maintained separately for protein and ionic liquid by use of a velocity rescaling thermostat^[54] with a time constant (τ) of 0.1 ps. Pressure coupling was done with a Berendsen barostat^[55] and use of a 1 bar reference pressure, isotropic pressure coupling, and a time constant of 4 ps. Electrostatic interactions were calculated by particle mesh Ewald (PME) summation,^[56,57] and twin range cutoffs of 1.0 and 1.4 nm were applied for computing the van der Waals interactions. $[\text{C}_2\text{mim}]^+$ and $[\text{OAc}]^-$ were parameterized by use of the automated topology builder^[58] and employment of a HF/STO-3G level of theory to compute the partial charges, which were subsequently scaled by a factor of 0.8 to model IL fluidity correctly.^[59–62] By using conformer 1 of the μ -conotoxin SIIIA NMR structure as a starting point,^[30] two simulation systems were constructed. In the first system, μ -SIIIA was placed inside a cubic box with an edge length of 5.14 nm and solvated with 485 $[\text{C}_2\text{mim}]^+$ and 488 $[\text{OAc}]^-$ molecules. In the second system, μ -SIIIA was converted into an extended conformation by using PyMol 1.4,^[63] placed inside a cubic box with an edge length of 9.21 nm, and solvated with 2490 $[\text{C}_2\text{mim}]^+$ and 2492 $[\text{OAc}]^-$ molecules. After steepest descent energy minimization, both systems were simulated for 100 ns.

Simulation analysis: To identify hotspots of IL- μ -SIIIA interaction, we computed the average number of IL contacts within a distance of 2.5 Å or less for each μ -SIIIA residue by use of the GROMACS tool g mindist and QtIPLOT for visualization. Additionally, the contact frequencies were mapped color-coded onto μ -SIIIA simulation conformations displaying the smallest root mean square displacement from the simulation average structures. As in refs. [64]–[67], simulation average structures were determined by using an iterative scheme of calculating the average conformation and realigning the trajectory to that average structure before computing a new average structure. This procedure was repeated until the average structure stopped changing.

Instrumentation for peptide analytics

Peptide purification: The crude peptides were purified by semipreparative reversed-phase HPLC with a Shimadzu LC-8A system equipped with a C18 column (Knauer Eurospher 100, Berlin, Germany). The gradient elution system was 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile/water (9:1; eluent B). The peptides were eluted with a gradient of 0–50% eluent B in 120 min and a flow rate of 10 mL min⁻¹. The peaks were detected at 220 nm. Collected fractions were combined, freeze-dried, and stored at –20 °C. Purities of the peptides were confirmed by analytical reversed-phase HPLC with a Shimadzu LC-10AT chromatograph (Duisburg, Germany) equipped with a Vydac 218TP54 column (C18, 5 μm particle size, 300 Å pore size, 4.6 \times 25 mm). The peptides were eluted with a gradient of 0–40% eluent B in 40 min. The flow rate was 1 mL min⁻¹, eluent A was 0.1% TFA in water, and eluent B was 0.1% TFA in acetonitrile; detection was at 220 nm.

Amino acid analysis: The amino acid compositions of the peptides were verified by amino acid analysis with a LC 3000 system from Eppendorf-Biotronik (Hamburg, Germany). Hydrolysis was performed with HCl (6N) in sealed tubes at 110 °C for 24 h. The analyses obtained supported the expected quantitative results. Amino

acid analysis in combination with HPLC (calibration curve) was used for the determination of peptide concentrations in solution prior to electrophysiological experiments.

Ellman's test: The procedure used is based on the method described by Ellman and co-workers and was described earlier.^[32] In brief, the measurements were carried out with a UV-1800 UV/Vis spectrophotometer (Shimadzu, Duisburg, Germany). The reference solution containing 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB, 19.8 mg) was prepared with Tris-HCl buffer (0.2 M, 50 mL, pH 8). Samples to be analyzed were prepared in DTNB/buffer solution and measured relative to the reference solution in quartz cuvettes of 1 cm pathlength.

Mass spectrometry: The molecular weights of crude and purified peptides were confirmed by MALDI-TOF mass spectrometry with an autoflex Bruker mass spectrometer (Bruker Daltonics, Bremen, Germany) and use of either α -cyano-4-hydroxycinnamic acid or sinapinic acid as matrices. The reflector mode was used for measurements and Peptide Calibration Standard II for molar masses $< 6000 \text{ g mol}^{-1}$.

Electrophysiological assay: HEK 293 cells (CAMR, Porton Down, Salisbury, UK) were maintained in Dulbecco's Modified Eagle's Medium (DMEM, 45 %) and Ham's F12 Medium (45 %), supplemented with fetal calf serum (10 %) in an incubator (5 % CO_2) at 37 °C. Cells were co-transfected with a 3:1 ratio of the rat $\text{Na}_v1.4$ expression plasmid^[68] and a vector encoding the CD8 antigen^[69] by use of the Rotifect transfection kit (Roth, Karlsruhe, Germany). Dynabeads (Deutsche Dynal, Hamburg, Germany) were used for visual identification of individual transfected cells. Na^+ current was measured by applying the whole-cell configuration of the patch-clamp method to HEK 293 cells 24–48 h after transfection. The patch pipettes contained: NaCl (35 mM), CsF (105 mM), ethylene glycol bis(2-amino-ethylether)tetraacetic acid (EGTA; 10 mM), HEPES (pH 7.4 with CsOH, 10 mM). The bath solution contained: NaCl (150 mM), KCl (2 mM), CaCl_2 (1.5 mM), MgCl_2 (1 mM), HEPES (pH 7.4 with NaOH, 10 mM). As patch-clamp amplifier an EPC-10 operated by PatchMaster software (HEKA Elektronik, Lambrecht, Germany) was used. Series resistance was corrected electronically up to 85 %; recordings with series resistance greater than 5 M Ω were discarded. μ -SIIIA fractions were diluted in bath solution and stored at –20 °C until use. Peptides were locally applied with a glass pipette as described earlier.^[70] Blocking of $\text{Na}_v1.4$ currents by μ -SIIIA fractions was assayed with repetitive pulses to 0 mV from a holding potential of –120 mV. The repetition interval was 3 s, and μ -SIIIA concentrations were 10 μM for the buffer-oxidized peptide and 90 μM for IL-treated and IL-oxidized peptides. μ -SIIIA-induced blocking of $\text{Na}_v1.4$ was determined by analyzing the inhibition of the peak currents after peptide application for 500 s. All experiments were performed at room temperature (25 °C).

Enzymatic assay: The activity of trypsin (Merck) was measured with $\text{N}\alpha$ -benzoyl-DL-arginine-4-nitroanilide hydrochloride as substrate and use of a UV/Vis spectrophotometer (Thermo Fisher Multiskan GO). Substrate concentration was 0.214 mM in a buffer containing Tris-HCl (pH 7.4, 40 mM), CaCl_2 (20 mM), and DMSO (25 %, v/v).^[71] Trypsin was dissolved in HCl (1 mM) at a concentration of 0.5 mg mL^{-1} . All μ -SIIIA fractions were dissolved in the same buffer as used in the electrophysiological assay (bath solution; see above). For each measurement, substrate solution (100 μL) was mixed with buffer and/or μ -SIIIA solution (100 μL) on a microtiter plate. The plate was then incubated for 5 min at 37 °C. Trypsin solution (5 μL) was then added. After 10 s of shaking, the absorption was measured at 410 nm for 10 min, with one data point taken

every 10 s. All measurements were performed at room temperature (25 °C).

Acknowledgements

The authors would like to thank Marianne Engeser (University of Bonn), Barbara Kirchner (University of Bonn), and Ute Neugebauer (Jena University Hospital) for support and access to MS facilities and Raman spectrometry, respectively, as well as Toni Kühl and Ming Chen for technical support. Donations of ionic liquids from BASF SE (Germany) and Sigma-Aldrich Chemie GmbH (Switzerland) (to A.S.) are gratefully acknowledged. Financial support for this work was provided by the Deutsche Forschungsgemeinschaft within priority program "SPP 1191: Ionic Liquids" (to R.G., D.I., and A.S.) as well as the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen (to C.K.). R.G. and A.K. thank the project "Sustainable Chemical Synthesis (SusChemSys)" which is co-financed by the European Regional Development Fund (ERDF) and the state of North Rhine-Westphalia, Germany, under the Operational Program "Regional Competitiveness and Employment" 2007–2013.

Keywords: conotoxins • cysteine-rich • ionic liquids • oxidative folding • peptide synthesis

- [1] F. van Rantwijk, R. A. Sheldon, *Chem. Rev.* **2007**, *107*, 2757–2785.
- [2] Y. Yu, X. Lu, Q. Zhou, K. Dong, H. Yao, S. Zhang, *Chem. Eur. J.* **2008**, *14*, 11174–11182.
- [3] W. Miao, T. H. Chan, *Acc. Chem. Res.* **2006**, *39*, 897–908.
- [4] A. A. Tietze, P. Heimer, A. Stark, D. Imhof, *Molecules* **2012**, *17*, 4158–5185.
- [5] H. Zhao, S. V. Malhotra, *Aldrichimica Acta* **2002**, *35*, 75–83.
- [6] Q. Zhang, S. Zhang, Y. Deng, *Green Chem.* **2011**, *13*, 2619–2637.
- [7] S. Toma, R. Sebesta in *Ionic Liquids in Biotransformation and Organocatalysis: Solvents and Beyond* (Ed.: P. Dominguez de Maria), Wiley-VCH, Weinheim, **2012**, pp. 333–361.
- [8] H. Zhao, *J. Mol. Catal. B* **2005**, *37*, 16–25.
- [9] Z. Yang in *Ionic Liquids in Biotransformation and Organocatalysis: Solvents and Beyond* (Ed.: P. Dominguez de Maria), Wiley-VCH, Weinheim, **2012**, pp. 15–66.
- [10] P. Dominguez de Maria in *Ionic Liquids in Biotransformation and Organocatalysis: Solvents and Beyond* (Ed.: P. Dominguez de Maria), Wiley-VCH, Weinheim, **2012**, pp. 3–14.
- [11] R. D. Rogers, *Nature* **2007**, *447*, 917–918.
- [12] H.-M. Choi, I. Kwon, *Ind. Eng. Chem. Res.* **2011**, *50*, 2452–2454.
- [13] a) Y. Ni, *J. Biocatal. Biotransformation* **2012**, *1*, 1–2; b) L. Andreani, J. D. Rocha, *Braz. J. Chem.* **2012**, *29*, 1–13; c) M. Hayyan, F. Mjalli, M. A. Hashim, I. M. AlNashef, *Fuel Process. Tech.* **2010**, *91*, 116–120.
- [14] A. Miloslavina, E. Leipold, M. Kijas, A. Stark, S. H. Heinemann, D. Imhof, *J. Pept. Sci.* **2009**, *15*, 72–77.
- [15] A. Miloslavina, C. Ebert, D. Tietze, O. Ohlenschläger, C. Englert, M. Gorchach, D. Imhof, *Peptides* **2010**, *31*, 1292–1300.
- [16] M. Böhm, T. Kühl, K. Hards, R. Coch, C. Arkona, B. Schlott, T. Steinmetzer, D. Imhof, *ChemMedChem* **2012**, *7*, 326–333.
- [17] A. Walewska, M. M. Zhang, J. J. Skaliky, D. Yoshikami, B. M. Olivera, G. Bulaj, *Angew. Chem. Int. Ed.* **2009**, *48*, 2221–2224.
- [18] K. H. Gowd, V. Yarotsky, K. S. Elmslie, J. J. Skaliky, B. M. Olivera, G. Bulaj, *Biochemistry* **2010**, *49*, 2741–2752.
- [19] A. A. Tietze, F. Bordusa, R. Giernoth, D. Imhof, T. Lenzer, A. Maaß, C. Mrestani-Klaus, I. Neundorff, K. Oum, D. Reith, A. Stark, *ChemPhysChem*, **2013**, *14*, 4044–4064.
- [20] S. Bräutigam, D. Dennewald, M. Schürmann, J. Lutje-Spelberg, W.-R. Pitner, D. Weuster-Botz, *Enz. Microbial. Technol.* **2009**, *45*, 310–316.

- [21] a) H. Weingärtner, C. Cabrele, C. Herrmann, *Phys. Chem. Chem. Phys.* **2012**, *14*, 415–426; b) D. R. MacFarlane, R. Vijayaraghavan, H. N. Ha, A. Izgorodin, K. D. Weaver, G. D. Elliott, *Chem. Commun.* **2010**, *46*, 7703–7705.
- [22] P. Attri, P. Venkatesu, A. Kumar, N. Byrne, *Phys. Chem. Chem. Phys.* **2011**, *13*, 17023–17026.
- [23] M. Ekimova, D. Fröhlich, S. Stalke, T. Lenzer, K. Oum, *ChemPhysChem* **2012**, *13*, 1854–1859.
- [24] M. M. Waichigo, B. M. Hunter, T. L. Riechel, N. D. Danielson, *J. Liq. Chromatogr. Rel. Techn.* **2007**, *30*, 165–184.
- [25] S. Grossman, N. D. Danielson, *J. Chromatogr. A* **2009**, *1216*, 3578–3586.
- [26] M.-M. Huang, H. Weingärtner, *ChemPhysChem* **2008**, *9*, 2172–2173.
- [27] M.-M. Huang, Y. Jiang, P. Sasisanker, G. W. Driver, H. Weingärtner, *J. Chem. Eng. Data* **2011**, *56*, 1494–1499.
- [28] L. W. Guddat, J. A. Martin, L. Shan, A. B. Edmundson, W. R. Gray, *Biochemistry* **1996**, *35*, 11329–11335.
- [29] C. Z. Wang, H. Zhang, H. Jiang, W. Y. Lu, Z. Q. Zhao, C. W. Chi, *Toxicol.* **2006**, *47*, 122–132.
- [30] S. Yao, M. M. Zhang, D. Yoshikami, L. Azam, B. M. Olivera, G. Bulaj, R. S. Norton, *Biochemistry* **2008**, *47*, 10940–10949.
- [31] K. J. Shon, B. M. Olivera, M. Watkins, R. B. Jacobsen, W. R. Gray, C. Z. Floresca, L. J. Cruz, D. R. Hillyard, A. Brink, H. Terlau, D. Yoshikami, *J. Neurosci.* **1998**, *18*, 4473–4481.
- [32] A. A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Kühl, A. Mischo, G. Buntkowsky, M. Görlach, S. H. Heinemann, D. Imhof, *Angew. Chem. Int. Ed.* **2012**, *51*, 4058–4061.
- [33] L. Volpon, H. Lamthanh, J. Barbier, N. Gilles, J. Molgo, A. Menez, J. M. Lancelin, *J. Biol. Chem.* **2004**, *279*, 21356–21366.
- [34] R. DeLa Cruz, F. G. Whitby, O. Buczek, G. Bulaj, *J. Pept. Res.* **2003**, *61*, 202–212.
- [35] S. Zorn, E. Leipold, A. Hansel, G. Bulaj, B. M. Olivera, H. Terlau, S. H. Heinemann, *FEBS Lett.* **2006**, *580*, 1360–1364.
- [36] a) R. Lungwitz, S. Spange, *New J. Chem.* **2008**, *32*, 392–394; b) R. Lungwitz, TU Chemnitz, private communication, **2011**.
- [37] M. Sellin, B. Ondruschka, A. Stark, *ACS Symp. Series* **2010**, *1033*, 121–135.
- [38] A. Stark, *Top. Curr. Chem.* **2010**, *290*, 41–81.
- [39] H. Zhao, C. L. Jones, J. V. Cowins, *Green Chem.* **2009**, *11*, 1128–1138.
- [40] K. Fujita, D. R. MacFarlane, M. Forsyth, M. Yoshizawa-Fujita, K. Murata, N. Nakamura, H. Ohno, *Biomacromolecules* **2007**, *8*, 2080–2086.
- [41] a) S. Oppermann, F. Stein, U. Kragl, *Appl. Microbiol. Biotechnol.* **2011**, *89*, 493–499; b) S. Dreyer, P. Salim, U. Kragl, *Biochem. Eng. J.* **2009**, *46*, 176–185; c) S. Dreyer, U. Kragl, *Biotechnol. Bioeng.* **2008**, *99*, 1416–1424.
- [42] J. V. Rodrigues, V. Prosinecki, I. Marrucho, L. P. N. Rebelo, C. M. Gomes, *Phys. Chem. Chem. Phys.* **2011**, *13*, 13614–13616.
- [43] S. Schlücker, C. Liang, K. R. Strehle, J. J. DiGiovanna, K. H. Kraemer, I. W. Levin, *Biopolymers* **2006**, *82*, 615–622.
- [44] A. Kumar, P. Venkatesu, *Chem. Rev.* **2012**, *112*, 4283–4307.
- [45] M. Brehm, H. Weber, A. S. Pensado, A. Stark, B. Kirchner, *Phys. Chem. Chem. Phys.* **2012**, *14*, 5030–5044.
- [46] T. Takekiyo, K. Yamazaki, E. Yamaguchi, H. Abe, Y. Yoshimura, *J. Phys. Chem. B* **2012**, *116*, 11092–11097.
- [47] A. Stark, P. Behrend, O. Braun, A. Müller, J. Ranke, B. Ondruschka, B. Jasstorff, *Green Chem.* **2008**, *10*, 1152–1161.
- [48] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Com. Phys. Comm.* **1995**, *91*, 43–56.
- [49] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- [50] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, W. F. van Gunsteren, *Eur. Biophys. J.* **2011**, *40*, 843–856.
- [51] C. Margreitter, D. Petrov, B. Zagrovic, *Nucleic Acids Res.* **2013**, *41*, W422–W426.
- [52] D. Petrov, C. Margreitter, M. Grandits, C. Oostenbrink, B. Zagrovic, *Plos Comput. Biol.* **2013**, *9*, e1003154.
- [53] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [54] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014 101 1–7.
- [55] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [56] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- [57] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- [58] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, A. E. Mark, *J. Chem. Theory Comput.* **2011**, *7*, 4026–4037.
- [59] M. Kohagen, M. Brehm, Y. Lingscheid, R. W. Giernoth, J. R. Sangoro, F. Kremer, S. Naumov, C. Iacob, J. Kärger, R. Valiullin, B. Kirchner, *J. Phys. Chem. B* **2011**, *115*, 15280–15288.
- [60] M. Kohagen, M. Brehm, J. Thar, W. Zhao, F. Müller-Plathe, B. Kirchner, *J. Phys. Chem. B* **2011**, *115*, 693–702.
- [61] B. Kirchner, *Top. Curr. Chem.* **2010**, *290*, 213–262.
- [62] S. Koßmann, J. Thar, B. Kirchner, P. A. Hunt, T. Welton, *J. Chem. Phys.* **2006**, *124*, 174506.
- [63] W. L. DeLano, *The PyMol Molecular Graphics System*, DeLano Scientific, San Carlos, CA, USA, **2002**.
- [64] N. Fischer, C. Kandt, *Proteins* **2011**, *79*, 2871–2885.
- [65] N. Fischer, C. Kandt, *Biochim. Biophys. Acta Biomembranes* **2013**, *1828*, 632–641.
- [66] D. C. Koch, M. Raunest, T. Harder, C. Kandt, *Biochemistry* **2013**, *52*, 178–187.
- [67] M. Raunest, C. Kandt, *Biochemistry* **2012**, *51*, 1719–1729.
- [68] J. S. Trimmer, S. S. Cooperman, S. A. Tomiko, J. Zhou, S. M. Crean, M. B. Boyle, R. G. Kalen, Z. Sheng, R. L. Barchi, F. J. Sigworth, R. H. Goodman, W. S. Agnew, G. Mandel, *Neuron* **1989**, *3*, 33–49.
- [69] M. E. Jurman, L. M. Boland, Y. Liu, G. Yellen, *Biotechnique* **1994**, *17*, 876–881.
- [70] H. Chen, D. Gordon, S. H. Heinemann, *Pflügers Arch.* **2000**, *439*, 423–432.
- [71] L. C. Correia, A. C. Bocewicz, S. A. Esteves, M. G. Pontes, L. M. Versieux, S. M. R. Teixeira, M. M. Santoro, M. P. Bemquerer, *J. Chem. Educ.* **2001**, *78*, 1535–1537.

Received: July 9, 2014

Published online on November 5, 2014

Epilogue

For the conotoxins tested, [C₂mim][OAc] turned out to be the IL of choice regarding the quality and yield of the oxidation product. A water content of > 3-5 % (v/v) disfavoured product formation or completely prevented the oxidation reaction. Increase in temperature to 60 or 80 °C increased the reaction speed and product yield, while for protic ILs the peptide was degraded after longer reaction time. Large scale production of μ -SIIIA in [C₂mim][OAc] revealed several fully oxidized species. These peptide forms were analyzed with HPLC, LC-ESI-MS/MS, and electrophysiological experiments revealing a lack in inhibitory activity at Na_v1.4-channel. There are two hypotheses about why the conotoxin lost its biological activity when produced in an IL. On the one hand, the IL-oxidation could result in different isomers of non-native cysteine connectivity on the other hand the IL ions might adhere to the peptide even after purification compromising the interaction with the biological target. The natively folded cysteine bonding, i.e. C1-C4, C2-C5 and C3-C6, was identified in the IL-oxidized peptide using LC-ESI-MS/MS after partial reduction and alkylation, thus excluding the possibility of misfolding. Computational molecular dynamics simulations of the peptide NMR structure in an IL environment revealed that in particular acetate anions preferably associate with the side chains of basic residues which are responsible for the ion channel block, thus hampering the toxin-channel interaction. Surprisingly, the adherence of ions to the peptide did not affect proteolytic degradation by trypsin, indicating that the effect at a biological target (e.g. enzyme vs. ion channel) is depending on the sensitivity of a specific ligand-protein interaction. In addition, buffer-oxidized, bioactive μ -SIIIA stored in IL also resulted in a loss of bioactivity providing further evidence for the impairment by adhering IL ions.

4.4 Molecular interaction of δ -conoptide EVIA with voltage-gated Na⁺ channels

Full Paper

Authors

Daniel Tietze, Enrico Leipold, Pascal Heimer, Miriam Böhm, Wadim Winschel, Diana Imhof, Stefan H. Heinemann, Alesia A. Tietze

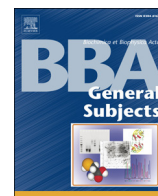
Biochimica et Biophysica Acta (BBA) - General Subjects, 1860, 9, 2016, 2053-2063.

doi: 10.1016/j.bbagen.2016.06.013

Copyright © 2016 Elsevier B.V. or its licensors or contributors. ScienceDirect[®] is a registered trademark of Elsevier B.V.

Preface

Identification of the binding site is essential to understand the mechanism of action of conotoxin binding to VGSCs.⁹⁷ It is essential to improve our understanding of possible side effects of such toxins, but respective studies are rare. In this report, a combination of molecular modelling and electrophysiology has been used to identify the binding site of δ -EVIA. This conotoxin was utilized herein to characterize the binding epitope for other voltage-sensor toxins such as δ -SVIE and scorpion α -toxins to support investigations on their selectivity for VGSCs.



Molecular interaction of δ -conopeptide EVIA with voltage-gated Na^+ channels

Daniel Tietze^{a,1}, Enrico Leipold^{b,1}, Pascal Heimer^c, Miriam Böhm^c, Wadim Winschel^a, Diana Imhof^c, Stefan H. Heinemann^b, Alesia A. Tietze^{d,*}

^a Eduard-Zintl-Institute of Inorganic and Physical Chemistry, Technische Universität Darmstadt, Alarich-Weiss-Str. 4, D-64287 Darmstadt, Germany

^b Department of Biophysics, Center for Molecular Biomedicine, Friedrich Schiller University Jena and Jena University Hospital, Hans-Knöll-Str. 2, D-07745 Jena, Germany

^c Pharmaceutical Chemistry I, Pharmaceutical Institute, University of Bonn, Brühler Str. 7, D-53119 Bonn, Germany

^d Clemens-Schöpf-Institute of Organic Chemistry and Biochemistry, Technische Universität Darmstadt, Alarich-Weiss-Str. 4, D-64287 Darmstadt, Germany

ARTICLE INFO

Article history:

Received 7 March 2016

Received in revised form 2 June 2016

Accepted 12 June 2016

Available online 15 June 2016

Keywords:

Conotoxin

δ -EVIA

NMR structural analysis

Voltage-sensor toxin

Molecular dynamics simulations

ABSTRACT

Background: For a large number of conopeptides basic knowledge related to structure-activity relationships is unavailable although such information is indispensable with respect to drug development and their use as drug leads. **Methods:** A combined experimental and theoretical approach employing electrophysiology and molecular modeling was applied for identifying the conopeptide δ -EVIA binding site at voltage-gated Na^+ channels and to gain insight into the toxin's mode of action.

Results: Conopeptide δ -EVIA was synthesized and its structure was re-determined by NMR spectroscopy for molecular docking studies. Molecular docking and molecular dynamics simulation studies were performed involving the domain IV voltage sensor in a resting conformation and part of the domain I S5 transmembrane segment. Molecular modeling was stimulated by functional studies, which demonstrated the importance of domains I and IV of the neuronal $\text{Na}_v1.7$ channel for toxin action.

Conclusions: δ -EVIA shares its binding epitope with other voltage-sensor toxins, such as the conotoxin δ -SVIE and various scorpion α -toxins. In contrast to previous *in silico* toxin binding studies, we present here *in silico* binding studies of a voltage-sensor toxin including the entire toxin binding site comprising the resting domain IV voltage sensor and S5 of domain I.

General significance: The prototypical voltage-sensor toxin δ -EVIA is suited for the elucidation of its binding epitope; in-depth analysis of its interaction with the channel target yields information on the mode of action and might also help to unravel the mechanism of voltage-dependent channel gating and coupling of activation and inactivation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The venoms of marine cone snails contain a spectrum of peptides (conotoxins) that typically act on ion channels and receptors and thus serve the snails for defense and prey capture [1–3]. Because of their specific target interaction, some conotoxins are of medicinal value; for example, the ω -conotoxin MVIIA from *Conus magus* (trade name:

Prialt®), which inhibits voltage-gated N-type Ca^{2+} channels in dorsal root ganglia neurons, is used as a potent analgesic [4–7]. Given that every cone snail species produces several hundreds of conopeptides, a plethora of biologically interesting substances is yet to be discovered [8,9]. A prerequisite for using such peptides for drug development, however, is a firm understanding of their structure and their molecular mode of action.

Most conotoxins harbor two or more cysteine residues; disulfide bridges stabilize their structure and determine their overall fold. Accordingly, conotoxins are classified into superfamilies according to their signal sequences, whereas disulfide bond patterns determine the cysteine framework [10]. Peptides of the M-superfamily that form three disulfide bridges have attracted considerable attention because they specifically inhibit voltage-gated ion channels. For example, μ -conotoxins interact with voltage-gated Na^+ (Na_v) channels by occluding the pore *via* binding in the extracellularly accessible channel vestibule [11]. This relatively simple interaction mechanism and the

Abbreviations: COSY, correlation spectroscopy; DG, distance geometry; DIEA, *N,N*-diisopropylethylamine; Fmoc, 9-fluorenylmethoxycarbonyl; GSH, glutathione reduced; GSSG, glutathione oxidized; HBTU, *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium-hexafluorophosphate; HSQC, heteronuclear single quantum coherence; NOESY, nuclear Overhauser enhancement spectroscopy; r.m.s.d., root mean square deviation; r.m.s.f., root mean square fluctuation; TFA, trifluoroacetic acid; TOCSY, total correlation spectroscopy.

* Corresponding author.

E-mail address: a.tietze@biochemie-tud.de (A.A. Tietze).

¹ Authors contributed equally.

perspective to develop μ -conotoxins as drugs to suppress neuronal signaling and therefore to be applied as analgesics has fueled research in this direction [12].

Conotoxins of the O-superfamily also contain three disulfide bridges, but they are much more diverse. Besides the already mentioned ω -conotoxins, which target voltage-gated Ca^{2+} channels, this superfamily comprises μO - and δ -conotoxins, both acting on Na_v channels [13]. Their mode of action, however, is very different from that of μ -conotoxins. While μ -conotoxins are “pore blockers”, μO -conotoxins and δ -conotoxins are so-called gating modifiers or voltage-sensor toxins, i.e. they interfere with parts of the Na_v channel proteins that are responsible for sensing changes in the transmembrane electric field (voltage sensors) and to gate the channel open. Although the crystal structure of the bacterial Na_vAB channel [14] has been solved, high-resolution structural data of mammalian Na_v channels are still lacking and homology considerations to Na_vAB and voltage-gated K^+ channels and numerous mutagenesis and functional studies have established the insight that Na_v channels consists of 4 homologous domains (I–IV), each of which with 6 transmembrane segments (S1–S6); S1–S4 form voltage sensors that move according to the electric field, while S5, S6 and the connecting linker form the channel gate and the central Na^+ -selective pore, respectively [15]. μO -conotoxins thus decrease Na_v channel activity by interfering with the voltage sensor of domain II and the S5-pore loop of domain III [13,16]. δ -Conotoxins, however, typically increase Na_v channel activity by interfering with rapid channel inactivation, and mutations in the S3–S4 linker of domain IV have a strong impact on this activity as demonstrated for δ -SVIE originating from *Conus striatus* [17]. Regarding the functional modulation of Na_v channels, δ -conotoxins share similarities with scorpion α -toxins, most likely having highly overlapping binding sites (Na_v channel toxin site 6), even though they show no sequence similarities [13,17].

Despite this intriguing activity, δ -conotoxins have attracted much less attention than other conopeptides probably because of their high hydrophobicity and limited availability. Besides synthetic δ -SVIE [17], δ -EVIA from *Conus ermineus* [18] and δ -CnVID from *Conus consors* [19] have been studied in more detail. In both of the latter cases, the toxin showed a remarkable preference for neuronal Na_v channels, while those from skeletal muscle ($\text{Na}_v1.4$) and the heart muscle ($\text{Na}_v1.5$) were largely unaffected by the toxins [18,19].

Given the advancement in peptide synthesis, novel structural data of voltage-gated ion channels, and progress in *in silico* description of peptide–protein interactions, we here aimed at providing insight into the molecular mechanism by which δ -EVIA interacts with mammalian Na_v channels and which aspects of the channel protein may be responsible for the toxin's subtype specificity.

δ -EVIA contains 32 amino acids with 4-hydroxyproline at position 6 and a C-terminal leucine amide. The six cysteines of the sequence are at positions 3, 10, 20, 21, 25, and 29 (cysteine pattern C–C–CC–C–C corresponding to the cysteine framework VI) with a connectivity as in representatives of μ - and ω -conotoxins: C1–4, C2–5, and C3–6 [18,20]. The structure of δ -EVIA was earlier reported and shown to be characterized by a 1:1 *cis/trans* isomerism of the Leu^{12} – Pro^{13} peptide bond within loop 2 (residues 11–19) of the disulfide-bridged sequence [21]. Both isomers were separately analyzed (PDB ID: 1G1P and 1G1Z) and, in addition, the exclusive occurrence of a *trans* leucyl–proline bond was demonstrated for the P13A analog, which was also synthesized by the same group. Interestingly, the latter was found to be only two-times less active than the parent peptide [21]. Another study reported on a peptidomimetic *cis* Pro–amide bond loop–analog of δ -EVIA; the aforementioned *cis/trans* peptide bond was exchanged using the pseudoproline approach involving $\text{Cys}(\Psi^{\text{Me,Me}}\text{Pro})^{13}$ that resulted in an analog with a fixed *cis* conformation of the respective bond in loop 2 [22].

Here we used a combined experimental and theoretical approach employing electrophysiology and molecular modeling in order to identify the δ -EVIA binding site and to gain insight into its mode of action. To ascertain functional assays and modeling to be based on the same

material and information, δ -EVIA was synthesized and its structure was re-determined by NMR spectroscopy. While functional evaluation revealed the importance of domains I and IV of the neuronal $\text{Na}_v1.7$ channel for toxin action, molecular docking and molecular dynamics simulation studies involving the domain IV voltage sensor in a resting conformation and part of the domain I S5 element revealed insight regarding how this voltage-sensor toxin specifically interferes with the mobility of the domain IV voltage sensor and thus delays channel inactivation.

2. Materials and methods

2.1. Peptide synthesis and chemical characterization

The linear peptide of δ -EVIA (sequence: DDCKOYGFCSLPILKNGLCSS GACVGVCADL– NH_2 ; O = 4-hydroxyproline) was synthesized and purified as described earlier [23]. Oxidation was carried out using an already established oxidative self-folding approach [11]. After the reaction was completed, the solvent was evaporated by freeze-drying, and concentrated aliquots were separated by RP-HPLC (see Supporting information). The molecular weight of crude and purified peptide was confirmed by mass spectrometry (see Supporting information).

Elucidation of the disulfide bonds of buffer-oxidized δ -EVIA was performed by LC-ESI MS/MS analysis. Therefore, an established protocol of partial reduction combined with carbamidomethylation was applied [24]. After sequencing of the linear, reduced peptide as control (Supporting Fig. S3A), the buffer-oxidized peptide was analyzed and the following connectivity was confirmed: C3–C21, C10–C25, C20–C29, corresponding to the suggested fold reported earlier (Supporting information, Supporting Fig. S3B, C) [21]. These findings confirm the results obtained from NMR structural analysis.

2.2. NMR analysis and structure calculation

Solution NMR experiments were performed at 293 K on a Bruker Avance III 600 MHz spectrometer. The peptide sample was dissolved in 90% H_2O /10% D_2O using the freeze-dried solid compound prepared in buffer solution. Data were acquired and processed with Topspin (Bruker, Rheinstetten, Germany) and analyzed with CCPN [25]. The proton resonance assignment was performed using a combination of 2D [^1H , ^1H]-TOCSY (80 ms spinlock time) and [^1H , ^1H]-NOESY experiments. Distance constraints were extracted from a [^1H , ^1H]-NOESY spectrum acquired with 120 ms mixing time. Upper limit distance constraints were calibrated according to their intensity in the NOESY spectrum. The three disulfide bridges were used as additional restraints. Torsion angle constraints were obtained from proton chemical shift analysis using DANGLE [26]. Structure calculations and refinements [27] were performed with YASARA structure [28–31]. The 16% of structures with the lowest energy were selected to represent the NMR ensemble for *trans*- δ -EVIA and *cis*- δ -EVIA, respectively (Supporting Fig. S4). Structures were refined in water at pH 4.

2.3. Homology modeling, molecular dynamics simulations and docking studies

The YASARA molecular modeling program was used to build homology models of $\text{Na}_v1.7$ and $\text{Na}_v1.4$ channels and for molecular dynamics simulations [29,30]. The voltage-sensor domain IV (VS-DIV) in the resting-state conformation and the adjacent full domain I (DI) of the respective voltage sensor were modeled independently. The resting-state model of the voltage-sensor domain of the voltage-gated K^+ channel ($\text{K}_v1.2$) created by Pathak et al. was used as a template for our resting-state VS-DIV homology models [32]. For the homology model of DI possible templates were identified by YASARA by running PSI-BLAST iterations [33] on the basis of the amino acid sequence of $\text{Na}_v1.4$ to extract a position specific scoring matrix (PSSM) from

UniRef90, followed by searching the Protein Data Bank (PDB) for a match. The templates were ranked based on the alignment score and the structural quality obtained from the PDBFinder2 database [34].

For the individual homology models of both domains, sequence alignments of the respective channel domain against the template structure were created with PSI-BLAST [33] against UniRef90. Multiple models of VS-DIV and DI were built according to alternative alignments of the target and template protein sequence. Side chains were added with YASARA's implementation of SCWRL3 [35] and refined considering electrostatic and knowledge-based packing iterations as well as solvation effects. The model was subsequently subjected to a fully unrestrained energy minimization with explicit solvent molecules by simulated annealing employing the YASARA2 force field [36]. In addition, a hybrid model was built by combination of the best regions of the individual models. The respective models were rated according to a given quality Z-score and the best scoring model was used for the final channel construct. A more detailed description of YASARA's homology modeling protocol can be found online (<http://yasara.org/homologymodeling.htm>).

Structure refinements of the homology models were performed as unrestrained all-atom molecular dynamics simulation for 10 ns in a membrane environment (PEA) and explicit water using the PME method [37] to describe long-range electrostatics at a cut-off distance of 8 Å and the YASARA2 [29] force field at physiological conditions (0.9% NaCl, 298 K, pH 7.4 [31]). The Amber03 [38] force field was used for the 50 ns structural refinements of the δ -EVIA-channel complex as well as for the respective reference simulations of rNa_v1.4 and mNa_v1.7 channel models. Docking was performed through 24 independent runs using VINA [39] with default parameters. The setup was done with the YASARA molecular modeling program [30]. Analysis of hydrogen bonds was performed after energy minimization of the toxin-protein complex after 50 ns membrane simulation in explicit water, using the PME method [37] and a cut-off distance of 8 Å to describe long-range electrostatics and the YASARA2 [29] force field (0.9% NaCl, 298 K, pH 7.4 [31]).

In order to qualitatively estimate the individual contribution of D4 S1–S2, D4 S3–S4 and D1 S5 to δ -EVIA binding, binding energies of δ -EVIA to the channel model were calculated as the difference between the potential energy (including solvation effects) of the complex and the sum of the individual potential energies of the ligand-free receptor and the free ligand.

In silico mutation analysis was performed after energy minimization of the best scoring docking result. For each mutation step, the side chain orientation of the swapped residue was optimized [35] followed by an energy minimization of the surrounding environment while the rest of the protein was kept fixed. Binding energies of the mutated complex were calculated using the VINA [39] scoring function as well as calculating the difference between the potential energy (including solvation effects) of the complex and the sum of the individual potential energies of the ligand-free receptor and the free ligand.

Molecular graphics were created with YASARA [40] (www.yasara.org) and POVray (www.povray.org).

2.4. Generation of channel expression constructs

The wild-type channel constructs used in this study were the rat skeletal muscle channel Na_v1.4 [41] (scn4a, P15390) and the mouse peripheral nerve channel Na_v1.7 (scn9a, Q62205); accession numbers refer to the UniProt database. Chimeras between Na_v1.4 and Na_v1.7 were assembled by replacing nucleotides 1–1328 (7444), 1329–3120 (4744), 3121–3863 (4474), 3864–5520 (4447), 1–1328 and 3864–5520 (7447) or 1329–3863 (4774) of scn4a with the homologous sequences of scn9a using a PCR-based strategy. The resulting channel chimeras contained either domain I (7444), domain II (4744), domain III (4474), domain IV (4447), domains I and IV (7447), or domains II and III of Na_v1.7 in a Na_v1.4 background. In addition, the following mutants

in the background of Na_v1.4 with variations of the domain IV S3–S4 linker were generated: Na_v1.4-L1424M:A1425F:S1427A:D1328E:L1429M:Q1431E (here termed 4444-rs7), Na_v1.4-L1424M:A1425F:S1427A (here termed 4444-MFA), and Na_v1.4-Q1431E (here termed 4444-QE). All mutant channel constructs were verified by DNA sequencing. Plasmid DNA was isolated from *Escherichia coli* using the Pure-Yield plasmid purification kit (Promega GmbH, Mannheim, Germany). Individual residues of the channel proteins are referenced in 1-letter code.

2.5. Electrophysiological recordings

HEK 293 cells (CAMR, Porton Down, Salisbury, UK) were maintained in Dulbecco's Modified Eagle's Medium (DMEM, 45%) and Ham's F12 Medium (45%), supplemented with fetal calf serum (10%) in an incubator (5% CO₂) at 37 °C. Cells were co-transfected with a 3:1 ratio of a wild-type or chimeric channel expression plasmid and a vector encoding the CD8 antigen using the Rotifect transfection kit (Roth, Karlsruhe, Germany). Dynabeads (Deutsche Dynal, Hamburg, Germany) were used for visual identification of individual transfected cells. Na⁺ current was measured by applying the whole-cell configuration of the patch-clamp method to HEK 293 cells 24–48 h after transfection. The patch pipettes contained (in mM): 35 NaCl, 105 CsF, 10 EGTA (ethylene glycol bis(2-amino-ethylether)tetraacetic acid), 10 HEPES (pH 7.4 with CsOH). The bath solution contained (in mM): 150 NaCl, 2 KCl, 1.5 CaCl₂, 1 MgCl₂, 10 HEPES (pH 7.4 with NaOH). As patch-clamp amplifier an EPC-10 operated by PatchMaster software (HEKA Elektronik, Lambrecht, Germany) was used. Series resistance was corrected electronically up to 80%. δ -EVIA aliquots were diluted in batch solution supplemented with 2 mg/ml BSA and stored at –20 °C until use. δ -EVIA was applied locally with a glass pipette as described earlier [42]. The effect of δ -EVIA on wild-type and chimeric channels was assayed with repetitive test pulses to –20 mV from a holding potential of –120 mV. The repetition interval was 3 s and δ -EVIA concentration was always 150 μ M. Removal of fast channel inactivation by δ -EVIA was evaluated by analyzing the mean current level between 4.5 and 5 ms after start of the depolarization (I_{5ms}) relative to the peak current (I_{peak}). The ratio I_{5ms}/I_{peak} gives an estimate of the probability of the channels not to be inactivated after 5 ms. Channel activation (I–V relationship) was assayed by step depolarizations ranging from –80 to 60 mV in steps of 10 mV. Peak currents were fit according to a Hodgkin–Huxley formalism according to [42, Eq. 1] yielding the half-maximal gate activation voltage (V_m) and the corresponding slope factor (k_m). Steady-state inactivation was measured for 500-ms time intervals at voltages between –140 and –40 mV in steps of 10 mV. The resulting remaining current, measured at –20 mV, was described with a single Boltzmann function yielding the half-maximal inactivation voltage (V_h) and the corresponding slope factor (k_h) [42, Eq. 2].

All experiments were performed at 20 \pm 1 °C. All data are presented as means \pm SEM (n = number of independent experiments). P values refer to two-sided Student's t -tests with unequal variances.

3. Results and discussion

3.1. Synthesis and NMR structure determination of δ -EVIA

The linear EVIA precursor was synthesized according to a standard Fmoc-SPPS protocol [23] (Supporting Fig. S1A). Oxidative folding (Supporting Fig. S1B) and the elucidation of disulfide bond connectivity were assessed through MS/MS analysis and NMR spectroscopy. Although the three-dimensional structure of δ -EVIA was reported earlier [21], we here reproduced the structure (Fig. 1A–C) as a foundation for the subsequent *in silico* binding studies. The overall fold of δ -EVIA reflects the inhibitor cysteine knot (ICK, native fold) motif found in numerous conopeptides of the O, P and I superfamilies and other toxic and inhibitory peptides [21,43,44]. Compared to other δ -conotoxins, δ -EVIA

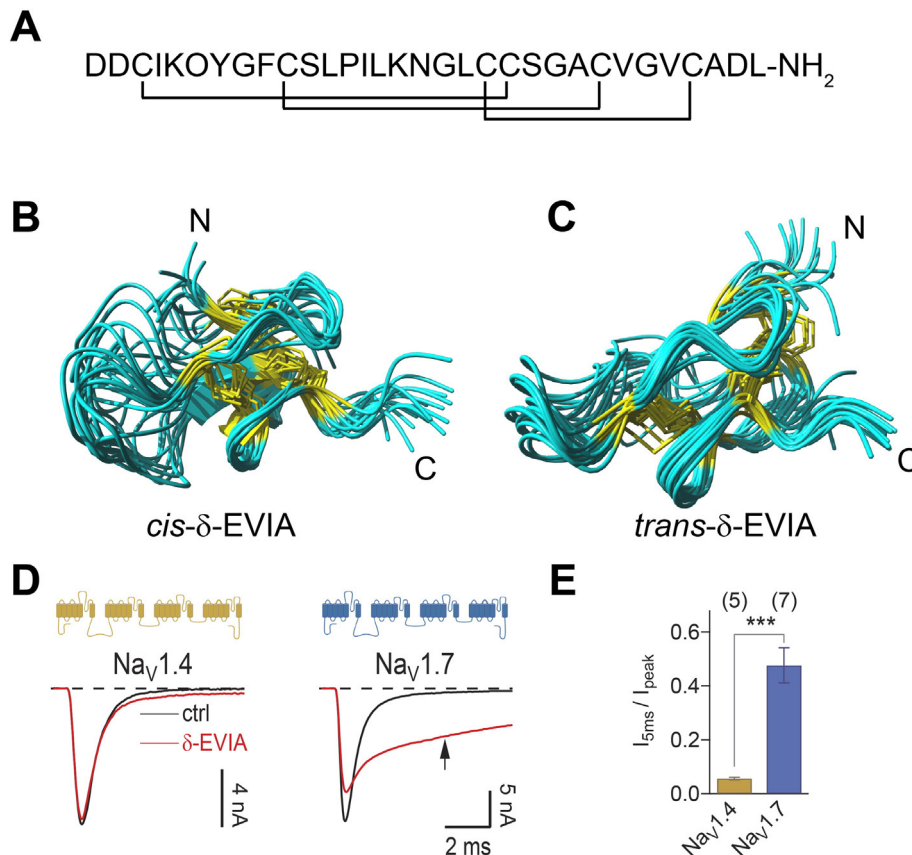


Fig. 1. Structure and activity of δ -EVIA. (A) δ -EVIA peptide sequence, including the location of disulfide bridges. (B, C) Solution NMR structure of *cis*- δ -EVIA (B) and *trans*- δ -EVIA (C) given as ensembles of 16 (*cis*) and 16 (*trans*) structures. Yellow indicates the position of cysteines. (D) Representative current traces obtained from HEK 293 cells heterologously expressing either $\text{Na}_v1.4$ or $\text{Na}_v1.7$ in response to a test voltage of -20 mV before (black) and after (red) application of $150 \mu\text{M}$ δ -EVIA. The arrow marks 5 ms after depolarization start, i.e. the time where loss of fast channel inactivation by δ -EVIA was analyzed. Cartoons on top of the current traces illustrate the membrane topology of the wild types $\text{Na}_v1.4$ (yellow) and $\text{Na}_v1.7$ (blue). (E) Mean of steady-state inactivation removal, expressed as $I_{5\text{ms}}/I_{\text{peak}}$, after application of $150 \mu\text{M}$ δ -EVIA; mean \pm SEM, n is given in parentheses; ***: $P < 0.001$. Electrophysiology studies were performed with purified oxidized δ -EVIA containing a conformational equilibrium of *cis* and *trans* conformations of the proline imide bond.

shows an extended loop-2 region [21]. This loop is of low conformational order. We also confirmed the coexistence of a *cis*- and *trans*-isomer of δ -EVIA in a 1:1 ratio, whose structures were solved independently in this study (Fig. 1B, C). The isomerism in δ -EVIA originates from a stable or slowly exchanging *cis* and *trans* peptide bond conformation between Leu12 and Pro13. The resonance assignments obtained are consistent with the earlier reported analysis [17]. The resulting structures for the δ -EVIA isomers have similar resolution as the previously reported structures (backbone r.m.s.d. δ -EVIA *trans* 1.62 Å and *cis* 1.91 Å compared to the published structures 1G1P (*trans*) 1.59 Å and 1G1Z (*cis*) 2.31 Å). The lowest energy structure for the δ -EVIA *trans* isomer deviates by 1.5 Å from the lowest energy structure of PDB ID 1G1P (Supporting Fig. S4). The lowest energy structure of *cis*- δ -EVIA deviates by 1.9 Å from that of the *cis*- δ -EVIA isomer available as PDB ID 1G1Z. The observed deviation for both isomers is mainly caused by the highly flexible loop 2 between residues Ser11 and Leu19. Overall, the herein reported structures for the δ -EVIA isomers are consistent with the structures reported earlier [17] and were thus used for the computational studies described below.

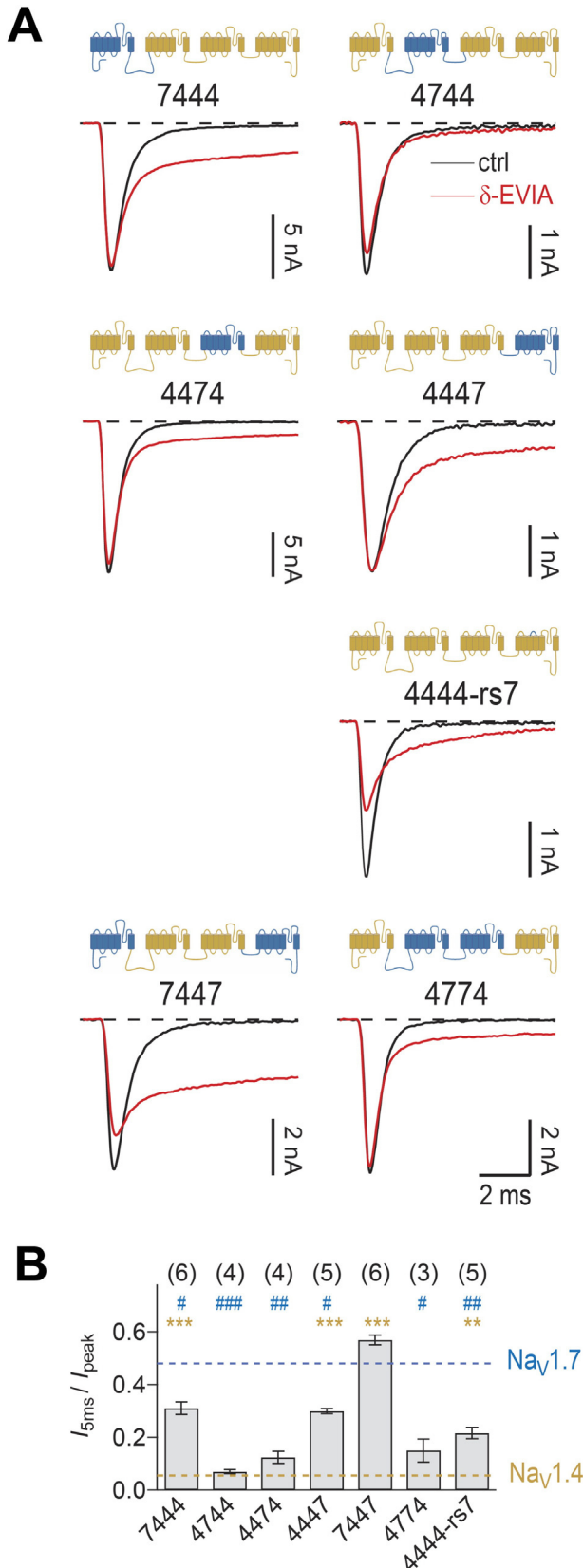
3.2. Functional properties of δ -EVIA

To infer about the biological activity of δ -EVIA, we assessed its effects on the inactivation of recombinant muscular $\text{Na}_v1.4$ and neuronal $\text{Na}_v1.7$ channels, expressed in HEK 293 cells. Consistent with a neuronal report where δ -EVIA was shown to interact preferentially with neuronal Na_v channels [18], $150 \mu\text{M}$ of δ -EVIA substantially slowed down inactivation of $\text{Na}_v1.7$, while it was almost inactive on $\text{Na}_v1.4$ (Fig. 1D, E).

This clear difference in sensitivity of $\text{Na}_v1.4$ and $\text{Na}_v1.7$ channels for δ -EVIA provided the basis for a systematic localization of channel domains that are important for δ -EVIA action. Thus, we constructed chimeric channels in which individual domains of $\text{Na}_v1.7$ were inserted into the background of the δ -EVIA insensitive channel $\text{Na}_v1.4$. As demonstrated in Fig. 2A, all domain chimeras gave rise to functional Na_v channels with fast inactivation characteristics as typical for $\text{Na}_v1.4$ and $\text{Na}_v1.7$. Moreover, channel chimeras displayed only minor alterations of their voltage dependence of activation (Supporting Fig. S12) and inactivation (Supporting Fig. S13), suggesting that the rearrangement of channel domains did not compromise normal channel function.

Functional assessment of such domain chimeras in voltage-clamp experiments revealed that δ -EVIA slowed down inactivation of the chimera that contains domain IV (4447) of $\text{Na}_v1.7$. This result is compatible with previous findings that δ -conotoxins exert their action by interfering with the voltage sensor of domain IV, specifically the S3/S4 linker [17], most likely because this voltage sensor is coupled to the intracellular inactivation gate [45–47]. Likewise, a chimera in which only the domain IV S3–S4 linker of $\text{Na}_v1.7$ was inserted in the background of $\text{Na}_v1.4$ (mutant 4444-rs7) exhibited a strong impact of δ -EVIA on channel inactivation, but neither chimera 4444-rs7 nor 4447 reached the δ -EVIA sensitivity of wild type $\text{Na}_v1.7$ (for statistics see Fig. 2B). However, δ -EVIA also markedly slowed down inactivation of chimera 7444, while the toxin had only a minor effect on channel inactivation when either domain II (4744) or III (4474) were from $\text{Na}_v1.7$ (Fig. 2A, B). Interestingly, δ -EVIA shares this domain-specificity with the structurally unrelated scorpion α -toxins, which also interfere with fast Na_v channel inactivation by interacting with domains I and IV [48,49]. Studies measuring

functional consequences of mixtures of scorpion α -toxins and δ -conotoxins [17] suggest that their receptor sites overlap, thus anticipating similar interaction epitopes.



Although the δ -EVIA-mediated loss of inactivation observed for chimeras 7444 and 4447 did not reach the level of Nav1.7 wild type, our data suggest that domains I and IV, but not domains II or III, possess the major determinants for the specific action of δ -EVIA. This became even more evident in constructs where domains I and IV were from one, and domains II and III from the other channel type. As shown in Fig. 2A and B, the full δ -EVIA effect, otherwise observed for Nav1.7 channels, was obtained for chimera 7447 ($I_{5ms}/I_{peak} = 0.48 \pm 0.07$ for Nav1.7 vs. 0.6 ± 0.02 for 7447; $P > 0.05$), whereas the δ -EVIA sensitivity of chimera 4774 rather resembled that of Nav1.4 ($I_{5ms}/I_{peak} = 0.06 \pm 0.01$ for Nav1.4 vs. 0.15 ± 0.04 for 4774; $P > 0.05$). These results clearly demonstrate that the interaction of δ -EVIA with Nav1.7 involves domains I and IV of Nav1.7. They also show that the corresponding domains of Nav1.4 clearly diminish the effect of δ -EVIA; however, this is not an all-or-nothing effect, thus anticipating that there might be only minor energetic differences in the molecular interaction of the δ -conotoxin with the corresponding receptor moieties of the channel proteins. These data thus favor a model in which the receptor site of δ -EVIA includes domains I and IV, where the interaction with the voltage sensor in domain IV mediates the slowdown of channel inactivation, while the interaction with domain I stabilizes the toxin–channel interaction.

According to the current traces shown in Fig. 2A, δ -EVIA not only affects the time course of inactivation but also reduces the maximal inward current in some cases. At a first glance this is unexpected because removal of fast inactivation alone is expected to yield an increase in peak current. We therefore measured the impact of δ -EVIA on the voltage dependence of activation (Supporting Fig. S12) and inactivation (Fig. 3A, Supporting Fig. S13) for all constructs. While only minor effects on channel activation were noticed, δ -EVIA partially had a strong left-shifting influence on the voltage dependence of inactivation. Since under standard conditions data were obtained from a holding potential of -120 mV, a left-shift of steady-state inactivation towards holding voltage is expected to reduce the number of available channels. As an operational data descriptor that includes the influence of the toxin on both, half-maximal inactivation voltage (V_h) and the voltage dependence of steady-state inactivation (k_h), we present the slope of steady-state inactivation between -140 and -120 mV in Fig. 3B. The impact of δ -EVIA on this slope for the constructs investigated shows a similar pattern as that of inactivation loss at -20 mV (Fig. 2B), indicating that the underlying molecular mechanisms are related or even identical.

Although we cannot provide unequivocal proof, the functional impact of δ -EVIA on Nav channels via interaction with the domain IV voltage sensor shares considerable similarity with the interaction of scorpion β -toxins [50] and μ O-conotoxins [51] with domain II voltage sensors. In both cases it seems that the toxin, once stabilized at the channel surface (presumably via interaction with domains I or III, respectively), can either retard the outward movement of the voltage sensor or trap it in an activated position. For δ -EVIA, retardation of voltage-sensor movement is expected to result in a slower time course of inactivation, while voltage-sensor trapping will promote closed-state inactivation and, hence, is expected to diminish the number of available channels and to reduce the peak current amplitude upon step depolarization. One consequence of such a scenario is that toxin interaction with the channels is strongly underestimated when the impact on the time course of inactivation is considered only.

Fig. 2. Functional evaluation of Nav channel domain chimeras. (A) Representative current traces obtained from HEK 293 cells heterologously expressing the indicated Nav chimeras in response to a test voltage of -20 mV before (black) and after (red) application of $150 \mu\text{M}$ δ -EVIA. The cartoons on top of the current traces illustrate the domain structure of the channel chimeras (yellow: Nav1.4, blue: Nav1.7). (B) Bar graph illustrating the removal of fast inactivation of the indicated Nav1.4/1.7 chimeras by $150 \mu\text{M}$ δ -EVIA. Dashed lines mark the ratios I_{5ms}/I_{peak} for wild-type channels Nav1.4 (yellow) and Nav1.7 (blue). Data are mean \pm SEM (n in parentheses); ***, $P < 0.001$, **, $P < 0.01$ tested against Nav1.4; #, $P < 0.05$, ##, $P < 0.01$, ###, $P < 0.001$ tested against Nav1.7.

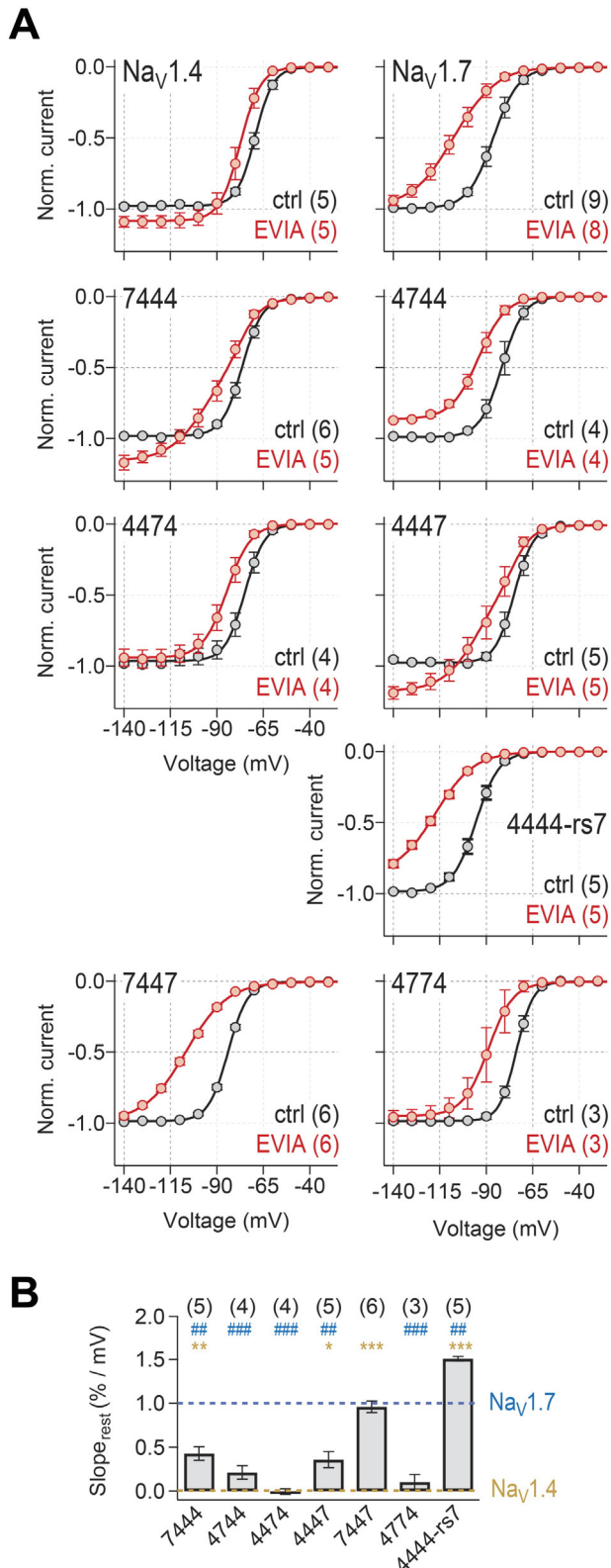


Fig. 3. Impact of δ -EVIA on steady-state inactivation. (A) Mean inward Na⁺ current at -20 mV measured after conditioning at various voltages for 500 ms before (black) and after (red) application of $150 \mu\text{M}$ δ -EVIA for the indicated constructs. Data are normalized to the values obtained at -120 mV (holding potential) of the control group. (B) Bar graph illustrating the slope of the steady-state inactivation between -140 and -120 mV of the indicated Na_v1.4/1.7 chimeras after application of $150 \mu\text{M}$ δ -EVIA. Dashed lines mark the ratios $I_{5\text{ms}}/I_{\text{peak}}$ for wild-type channels Na_v1.4 (yellow) and Na_v1.7 (blue). Data are mean \pm SEM (n in parentheses); ***, $P < 0.001$, **, $P < 0.01$, *, $P < 0.05$ tested against Na_v1.4; #, $P < 0.05$, ##, $P < 0.01$, ###, $P < 0.001$ tested against Na_v1.7.

3.3. Structure-function relationship derived from theoretical binding studies

3.3.1. Homology modeling

As mentioned earlier, the binding site of the scorpion α -toxin strongly overlaps with that of δ -conotoxins [17]. Hence, we focused on modeling the extracellular half of transmembrane segment S5 with part of the subsequent S5-SS1 loop and the S6-S6 loop of domain I, as well as the voltage sensor of domain IV because these regions were identified as the main binding sites for the much bigger scorpion α -toxin LqhII [52]. The aim of the subsequent experiment was an *in silico* evaluation of the interaction of δ -EVIA with this channel construct by performing theoretical binding studies employing molecular docking and molecular dynamics simulations.

Previous experiments with δ -SVIE suggest that the δ -conotoxins interact with the DIV voltage sensor in its resting state, and strong depolarizations, i.e. activation of the voltage sensors, diminish the toxin effect on channel inactivation [17]. Therefore, for *in silico* docking of δ -EVIA with Na_v channels, molecular models of the voltage sensors in the resting state are required. Although crystal structures of bacterial voltage-gated Na⁺ channels (Na_vRh PDB ID: 4DXW [53], Na_vAB-Ile217Cys, -wt PDB ID: 3RVY [14], 4EKW [54]) and voltage-gated K⁺ channels (VGPCs) (K_vAP PDB ID: 1ORQ [55]) are available in the RCSB database, the structures of mammalian Na_v channels, which are studied herein, are not yet described. Furthermore, the channel structures reported were crystallized in the open conformation and the conformation of the voltage-sensor domain represents the activated conformation with the S4 helix exposed to the external solution. Therefore, homology models of rat Na_v1.4 and mouse Na_v1.7 and the respective chimeras with the domain IV voltage sensor in the resting state had to be established. Instead of homology modeling of the full Na_v channels, we independently built models for domain I and the resting state of VS-DIV, which were subsequently aligned to the crystal structure of Na_vAB-Ile217Cys to bring both parts into the correct relative orientation (Supporting Fig. S6).

In order to obtain a first Na_v homology model (Fig. 4A), the respective domain parts of PDB ID 3rvy were replaced by the modeled parts and then refined in a 10-ns all-atom MD simulation in a membrane environment (PEA, phosphatidylethanolamine). This refined Na_v homology model was then used to build the remaining channel models by replacing the respective residues in order to maintain the backbone conformation for all channel models.

Guided by the results of the electrophysiological experiments, we constructed a set of four channel models representing Na_v1.4, Na_v1.7, as well as the chimeras 7444 and 4447 (Fig. 2A). The previously reported structure of K_v1.2 VGPC in the closed conformation [32], which revealed the voltage sensor in the resting state conformation, was used as a template to homology model the resting state conformation of the domain IV voltage sensor of Na_v1.4 and Na_v1.7. Furthermore, the obtained homology model was validated and aligned with the recently published cryo-EM structure of the rabbit voltage-gated Ca²⁺ channel Ca_v1.1 (PDB entry 3JBR) [56]. The relative orientation of VS-DIV S1-S4 and DI S5 and S6 is in good agreement in both structures (Supporting Fig. S15).

With the exception of the leading 9 residues in the DI S5-SS1 loop (Supporting Fig. S5), this channel part had a very poor structural quality and was designated as highly uncertain by the homology modeling software and therefore was excluded from our channel models. The four channel constructs were energy-minimized before they were used for the docking experiments. Docking was locally restricted to the protein area formerly highlighted as binding site for the scorpion α -toxin LqhII [52]. The best-scored docking run was subjected to an additional refinement step of 50 ns in a membrane environment (PEA); conformational stability during the entire simulation was verified (Supporting Fig. S9). Only the lowest energy conformation of the *cis* isomer of δ -EVIA was used for docking because it was found that it exhibits a higher bioactivity than the *trans* isomer [21].

3.3.2. *In silico* toxin binding studies

The docking results were scored according to the binding energy between the toxin and the channel construct (Supporting Table S2) and revealed the highest toxin binding energy for the δ -EVIA-Nav_v1.7 complex followed by the two δ -EVIA-Nav_v1.4 chimera complexes 4447 and 7444. The lowest toxin binding energy was observed for the δ -EVIA-Nav_v1.4 complex. Although the absolute energy values are not straightforwardly interpreted and should only be seen as a qualitative measure, the rank order of these interaction energies is in accord with the experimental observations (Figs. 1–3), which anticipate that domains I and IV of Nav_v1.7 make an additive contribution to the functional impact of δ -EVIA in channel inactivation. Images of the corresponding docking results are displayed in Fig. 4 and Supporting Fig. S10.

In all channel models evaluated, δ -EVIA occupies the cleft formed by the four transmembrane segments of the domain IV voltage-sensor domain. Together with the upper part of DI S5, they comprise the main δ -EVIA binding site (also see Supporting Fig. S11A). A closer look at the δ -EVIA-channel complexes indicates that the toxin adopts different orientations upon binding to models of Nav_v1.7, Nav_v1.4 (Fig. 4B, C), 4447 and 7444 (Supporting Fig. S10). As shown in Fig. 4B for the δ -EVIA-Nav_v1.7 complex, the flexible loop of δ -EVIA comprising residues Ser11 to Leu19 is located on top of the DIV S3–S4 loop, while the N-terminal Asp1 is buried deeply in the cleft between S2 and S3 of the voltage-sensor domain. In contrast, in the δ -EVIA-Nav_v1.4 complex (Fig. 4C) this loop points upwards and is aligned almost in parallel to the DIV S2 segment orienting the N terminus of δ -EVIA towards the upper end of S1. A similar orientation of δ -EVIA was observed when the toxin was bound to the 4447 channel model (Supporting Fig. S10A); the toxin is rotated by approximately 60° directing the N-terminal Asp1 towards S4. In the 7444 channel model, δ -EVIA adopts an almost identical orientation as observed for the δ -EVIA-Nav_v1.7 complex (Supporting Fig. S10B). However, the flexible loop of the toxin had slightly moved towards S3 compared to its orientation in the δ -EVIA-Nav_v1.7 complex.

Homology modeling of the channel constructs resulted in structures with global conformations not differing significantly. Thus, we can exclude differences in the starting conformations to be the reason for the different toxin orientations obtained through docking simulations.

In order to estimate how different domain parts of the channel protein contribute to toxin binding, we calculated individual binding energies (average binding energies calculated from the last 5 ns of the 50-ns refinement simulation, Supporting Fig. S9B) of δ -EVIA binding to DI S5, DIV S1–S2 and DIV S3–S4. Again, the binding energies calculated by VINA do not represent realistic values but will at least allow for a qualitative analysis of the docking results. As shown in Fig. 5A, all three channel parts contribute to the toxin binding, but their individual contribution differs among our four channel models. Interestingly, our data further indicate that DI S5 in Nav_v1.4 is more weakly bound to δ -EVIA compared to the other channel models.

We thus tried to identify potential key residues for toxin binding through an *in silico* cumulative mutagenesis study (Fig. 5B, C; Supporting Fig. S7A, B): residues that differ between the δ -EVIA-insensitive Nav_v1.4 and δ -EVIA-sensitive Nav_v1.7 channel and that belong to the δ -EVIA binding site were stepwise mutated towards the sensitive channel and *vice versa*. The change in binding energy for δ -EVIA binding was calculated for each mutation individually and is shown in Fig. 5B and C (for change in binding energy including solvation effects (see Supporting Fig. S7A, B)). Four single residues originating from Nav_v1.7 in the background of Nav_v1.4 were identified, which stand out with their positive impact on toxin binding over the other mutations (Fig. 5C). These mutations were Q1372D in the DIV S1–S2 linker, A1425F in the DIV S3–S4 linker and Q278H and V281F in the DI S5 segment. Most single mutations caused an increase in the δ -EVIA binding energy with the exception of D1369G in the DIV S1–S2 linker region and S1427A and D1428E in the DIV S3–S4 linker region, which caused a decrease in the toxin binding energy. The *in silico* mutation study of these residues (Fig. 5C) is consistent with the electrophysiological

experiments that identified the DIV S3–S4 linker as an important determinant (see chimera 4444-rs7 in Figs. 2 and 3). Moreover, our Nav_v1.4-EVIA structural model reveals that E1431 in the Nav_v1.4-Q1431E mutant remains hydrogen bound to Tyr7 of the toxin, while this interaction is not formed with the neutral Gln residue in the wild-type channel (Fig. 4D). Similarly, the model supports the rationale that the increased hydrophobicity in a Nav_v1.4-A1425F mutation leads to an increase in toxin binding affinity by embedding the C-terminal Leu32 side chain of the toxin (Fig. 4E, F). We thus also generated and assayed mutations in the background of Nav_v1.4 channels that reconstitute only parts of the DIV S3–S4 linker of Nav_v1.7. As documented in Supporting Fig. S14, mutations 4444-Q1431E and 4444-MFA (the latter representing the change A1425F) both make the channel more sensitive towards δ -EVIA.

Fig. 5B illustrates predictions for step-wise mutation of Nav_v1.7 to Nav_v1.4, suggesting a somewhat different relevance of individual residues. The difference might be explained by the different toxin orientations observed at Nav_v1.4 versus Nav_v1.7, which in turn would cause each single mutation step to have a different impact on the toxin binding affinity. A strong decrease in toxin affinity was found for mutations located in DIV S2 (especially F1532D and W1536N), in the DIV S3–S4 linker (especially F1582A, A1584S and E1585D) and a single mutation in DI S5 (H273Q). Interestingly, the residues W1536 in DI S2 as well as F1582 and A1584 in DIV S3–S4 linker are conserved among the δ -EVIA-sensitive Nav_v channel homologs (Supporting Fig. S11A) and presumably contribute to the subtype specificity of δ -EVIA.

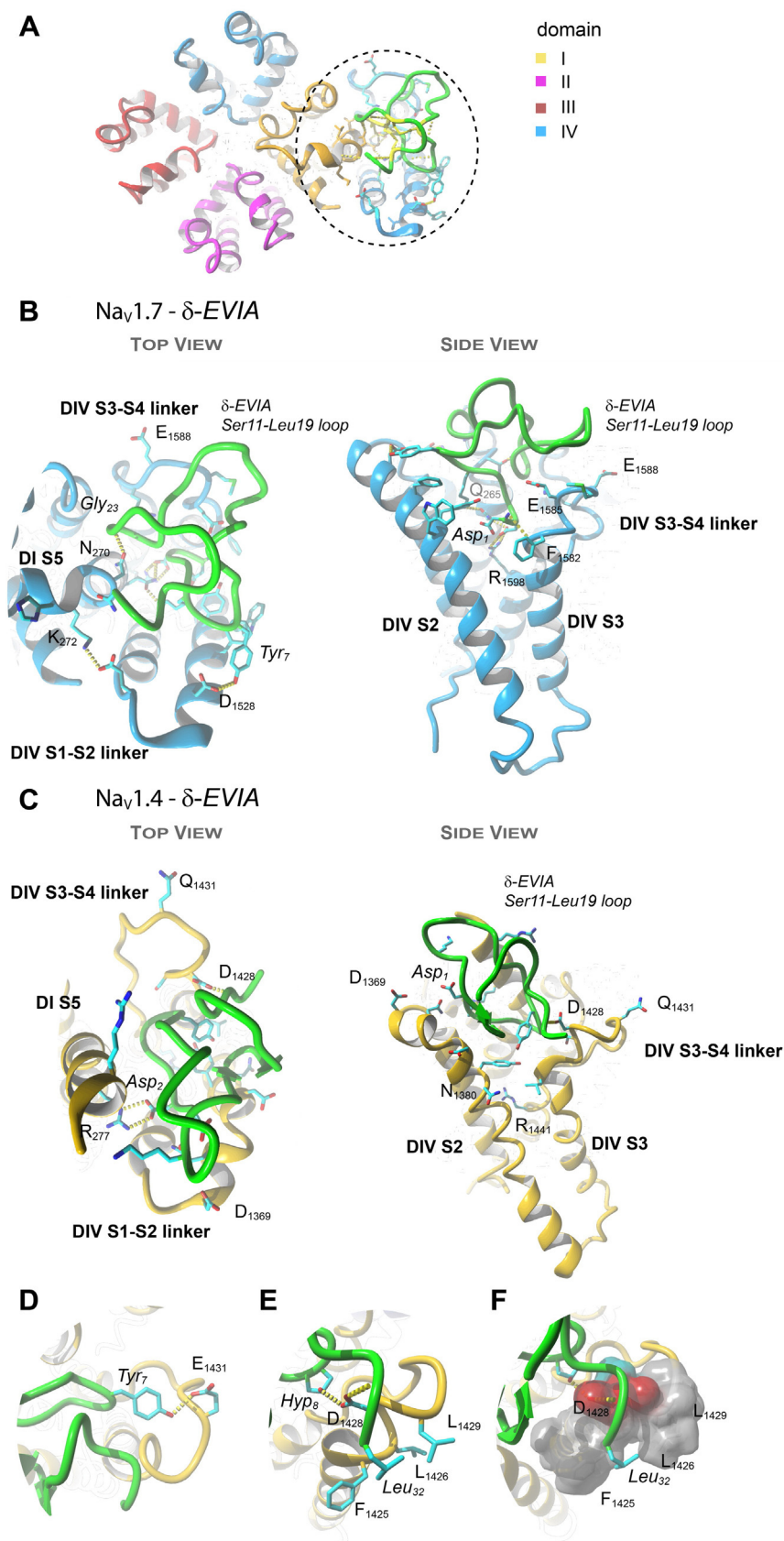
In order to identify further key interaction sites, the best scoring toxin-channel conformation from the docking runs was allowed to equilibrate over a 50-ns molecular dynamics simulation in a membrane environment. The structures were fully equilibrated reaching a stable conformation after about 30 ns according to the calculated r.m.s.d. (with respect to the initial structure) over the entire simulation run (Supporting Fig. S9A). Although toxin orientation differed upon binding to the channel models, we realized that δ -EVIA binds to residues in DIV S4 and DIV S3 in Nav_v1.7 and both toxin-sensitive chimeras (7444 and 4447) through salt bridge and/or hydrogen bond formation (Fig. 4B, C; Supporting Fig. S10). In particular, at Nav_v1.7 the side chain of Asp2 of δ -EVIA forms a salt bridge with the side chain of R1598 (Fig. 4B) in S4, while the N-terminal amine group of δ -EVIA forms a hydrogen bond to the protein backbone in S3 (M1581). Furthermore, Asp2 together with Lys5 of the toxin spans a hydrogen bond network to Q265 in S5 of domain I. Two additional hydrogen bonds are formed between the backbone of the toxin and the protein backbone of S5 (Fig. 4B).

In case of the 4447 chimera, the side chain of Lys5 of δ -EVIA forms a salt bridge to the side chain of E1585 in the S3–S4 linker region and a hydrogen bond to the protein backbone in S4 (T1595). Additionally, δ -EVIA (Asp1) forms a single salt bridge to S5 (N275). Interestingly, in the 7444 channel chimera, Asp1 and Asp2 form a salt bridge to R1441 (analogous to R1598 in Nav_v1.7; Supporting Fig. S10B, Supporting Fig. S11B) and span a hydrogen bond network involving the side chains of T1438 in S4, D1428 in the S3–S4 linker region and Y1379 in S2. However, no hydrogen bond or salt bridge to S5 was observed in this model. In contrast, DIV S4 is not involved in toxin binding when δ -EVIA binding is modeled at the insensitive Nav_v1.4 channel. Moreover, toxin binding involves mainly residues in S3 and S2 and only a single salt bridge between δ -EVIA and S5 was observed (Asp2 – R277; Supporting Fig. S10B). This particular toxin-channel interaction, especially to S4, might explain the observed delayed inactivation of Nav_v1.7 channels resulting from a hindered, thus slower outward movement of S4. The data moreover indicate that it seems to be sufficient to partly restore interactions between the toxin and the S3–S4 linker region in order to reintroduce a delayed channel inactivation by the toxin (Figs. 2 and 3, Supporting Fig. S14).

A more detailed analysis of the trajectories generated by the 50-ns refinement simulations of the best-scoring toxin-channel complexes provides interesting insights into the δ -EVIA binding and supports the idea of a toxin key interaction to S4 in order to efficiently slow down

fast inactivation of voltage-gated Na^+ channels. Therefore, we calculated average protein structures of the δ -EVIA-channel complex and used average per-residue fluctuations relative to the starting structure

(expressed as root mean square fluctuations, r.m.s.f.) as a measure of dynamic motion, which is often linked to intrinsic domain motion [57–59]. As depicted in Supporting Fig. S8, the local per-residue



flexibility is reduced upon δ -EVIA binding relative to the reference simulation with/without the toxin, especially in the linker region connecting the transmembrane segments of the voltage-sensor domain of Na_v1.7 and the toxin-sensitive chimera 4447.

4. Conclusions

The solution NMR structure of the conopeptide δ -EVIA determined herein is in good agreement with earlier reports [21] and was used for the *in silico* binding studies. δ -EVIA slowed down the inactivation of neuronal Na_v1.7 channels, while the effect was marginal for the Na_v1.4 from the skeletal muscle. Domain chimeras between these two Na_v channel isoforms revealed that domains I and IV of Na_v1.7 increased the sensitivity of the resulting constructs to δ -EVIA in an additive manner, while exchange of domains II and III was without noticeable functional consequences. Within domain IV, the linker between S3 and S4 is the most relevant structure for toxin action. Mechanistically, the impact of δ -EVIA on the voltage sensor of domain IV can be envisioned in two ways: interference with the voltage sensor in its deactivated position slows down its activation and, hence, also the time course of inactivation; interference in the activated position may lead to voltage-sensor trapping with the consequence of facilitated closed-state inactivation, which in effect will lead to a reduction of peak current amplitude.

Based on the functional insights, homology channel models were constructed, which contained the voltage sensor of domain IV in a resting conformation and the transmembrane segment S5 with part of domain I SS1. In a combination of *in silico* docking and molecular dynamics simulations, the interaction of δ -EVIA with the respective channel protein domains were assessed. The rank order to the estimated toxin binding energies were compatible with the experimental observation that domains I and IV of Na_v1.7 support the functional activity of δ -EVIA. *In silico* binding studies indicated that each transmembrane segment of the domain IV voltage sensor as well as S5 of domain I contribute to toxin binding. Although the orientation of the conotoxin on the channel differed between different channel models, the voltage-sensing S4 of domain IV remained its contact with the toxin at the δ -EVIA-sensitive Na_v channel constructs. The simulations further indicate that the outermost voltage-sensing arginine residues in DIV S4 take part in toxin binding. Recently, Chen and coworkers also reported for the scorpion α -toxins AahlI and LqhαIT that these arginine residues directly interact with the toxin [60,61]. This observation is compatible with the experimental finding that δ -conotoxins delay the outward movement of S4, thus resulting in a slowing down of fast channel inactivation. However, detailed molecular dynamics simulations of extended models and various binding configurations are required to obtain further insight into the sequence of individual molecular events linked to voltage-sensor activation with and without δ -conotoxin interference.

In summary, the conopeptide δ -EVIA from *Conus ermineus* exhibits a preference for neuronal voltage-gated Na⁺ channels. In a domain-swapping approach applied to neuronal Na_v1.7 and muscular Na_v1.4, domains I and IV were identified as the major determinants for that specificity. δ -EVIA therefore shares its binding epitope with other voltage-sensor toxins, such as the conotoxin δ -SVIE and scorpion α -toxins Lqh2, Lqh3, AahlI, and LqhαIT [17,52,60,61]. In contrast to previous *in silico* toxin binding studies [60,61], we present here an *in silico* binding study of a voltage-sensor toxin including the entire toxin

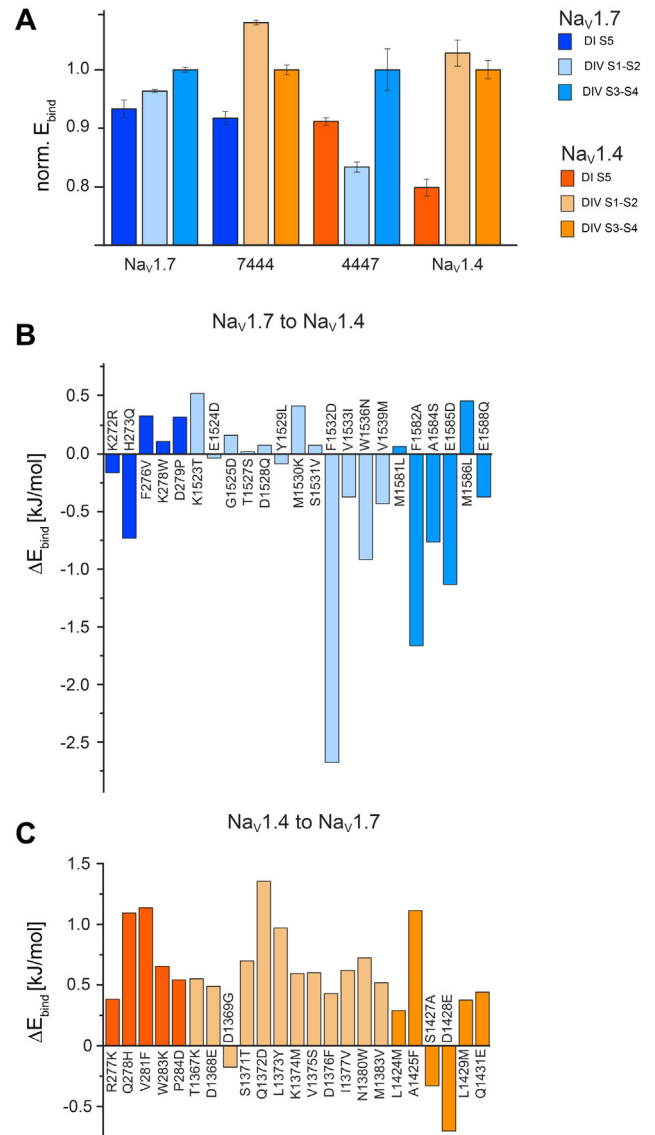


Fig. 5. (A) Individual contributions of DI S5, DIV S1–S2, and DIV S3–S4 to δ -EVIA binding calculated from the average binding energies through the last 5 ns of the refinement simulation. Binding energies are given as normalized values with respect to the binding energy of δ -EVIA to the DIV S3–S4 segment (set to one) for each channel construct. (B, C) *In silico* mutagenesis study: mutation scan for δ -EVIA binding at Na_v1.7 and Na_v1.4. For each mutation step, toxin binding energy was calculated individually and plotted as an increase or decrease with respect to the binding energy calculated from the wild-type channel.

binding site comprising the resting domain IV voltage-sensor and S5 of domain I.

In particular because of its small size, the prototypical voltage-sensor toxin δ -EVIA is suited for the elucidation of its binding epitope; in-depth analysis of its interaction with the channel target yields valuable information on the mode of action and might also help to unravel the mechanism of voltage-dependent channel gating and coupling of activation and inactivation. Future modeling attempts, however, will need to consider the non-conserved extracellular S5–S6 linkers of domain I – an intricate task because of their high structural flexibility.

Fig. 4. Na_v channel homology models and δ -EVIA docking results. (A) Top view of the Na_v1.7 channel model with docked δ -EVIA after 50 ns of refinement simulation. The voltage-sensor domains of domains I through III were excluded from the refinement simulation. Channel domains are color-coded; the toxin is shown in green. (B, C) Energy-minimized final structural models of *cis*- δ -EVIA docked to the binding sites of Na_v1.7 (B, close-up of dashed circle of panel A) and Na_v1.4 (C) after 50 ns refinement in a phospholipid environment in top (left) and side view (right). Hydrogen bonds are indicated as yellow dashed lines. R1598, which is the outermost arginine residue in S4 of domain IV, is involved in δ -EVIA binding at Na_v1.7, while the corresponding R1441 in Na_v1.4 is not. (D, E, F) Energy-minimized structures of the *in silico* mutagenesis study in the background of Na_v1.4. (D) Na_v1.4-Q1431E, (E) Na_v1.4-A1425F, (F) van der Waals surface representation of the hydrophobic region in the S3–S4 linker in Na_v1.4-A1425F illustrating embedding of the C-terminal toxin residue Leu32 (surface colors: gray, hydrophobic; red, acidic positively charged; cyan, polar uncharged).

Transparency Document

The Transparency Document associated with this article can be found, in online version.

Acknowledgements

We thank Dr. Volker Schmidts for NMR measurements and Prof. Dr. Christina M. Thiele (both TU Darmstadt) for scientific discussions. Financial support by the Liebig Fellowship of the Fonds der Chemischen Industrie im Verband der Chemischen Industrie e.V. (A.A.T.), the University of Bonn (to D.I.), the German Research Foundation (LE-2338/3-1 to E.L. and TI839/2-1 to D.T.), as well as the Fonds der Chemischen Industrie (to D.T.) are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbagen.2016.06.013>.

References

- [1] S. Dutertre, A.H. Jin, I. Vetter, B. Hamilton, K. Sunagar, V. Lavergne, V. Dutertre, B.G. Fry, A. Antunes, D.J. Venter, P.F. Alewood, R.J. Lewis, Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails, *Nat. Commun.* 5 (2014) 3521.
- [2] B.M. Olivera, J. Seger, M.P. Horvath, A.E. Fedosov, Prey-capture strategies of fish-hunting cone snails: behavior, neurobiology and evolution, *Brain Behav. Evol.* 86 (2015) 58–74.
- [3] B.M. Olivera, P.S. Corneli, M. Watkins, A. Fedosov, Biodiversity of cone snails and other venomous marine gastropods: evolutionary success through neuropharmacology, *Annu. Rev. Anim. Biosci.* 2 (2014) 487–513.
- [4] J.M. McIntosh, R.M. Jones, Cone venom—from accidental stings to deliberate injection, *Toxicol.* 39 (2001) 1447–1451.
- [5] C.I. Schroeder, C.J. Doering, G.W. Zamponi, R.J. Lewis, N-type calcium channel blockers: novel therapeutics for the treatment of pain, *Med. Chem.* 2 (2006) 535–543.
- [6] C.I. Schroeder, R.J. Lewis, ω -Conotoxins GVIA, MVIIA and CVID: SAR and clinical potential, *Mar. Drugs* 4 (2006) 193–214.
- [7] D.P. Wermeling, Ziconotide, an intrathecally administered N-type calcium channel antagonist for the treatment of chronic pain, *Pharmacotherapy* 25 (2005) 1084–1094.
- [8] J. Davis, A. Jones, R.J. Lewis, Remarkable inter- and intra-species complexity of conotoxins revealed by LC/MS, *Peptides* 30 (2009) 1222–1227.
- [9] S. Dutertre, A.H. Jin, Q. Kaas, A. Jones, P.F. Alewood, R.J. Lewis, Deep venomomics reveals the mechanism for expanded peptide diversity in cone snail venom, *Mol. Cell. Proteomics* 12 (2013) 312–329.
- [10] Q. Kaas, J.C. Westermann, D.J. Craik, Conopeptide characterization and classification: an analysis using ConoServer, *Toxicol.* 55 (2010) 1491–1509.
- [11] A.A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Kuhl, A. Mischo, G. Buntkowsky, M. Gorlach, S.H. Heinemann, D. Imhof, Structurally diverse μ -conotoxin PIIIA isomers block sodium channel Na(V) 1.4, *Angew. Chem. Int. Ed.* 51 (2012) 4058–4061.
- [12] R.S. Norton, μ -Conotoxins as leads in the development of new analgesics, *Molecules* 15 (2010) 2825–2844.
- [13] S.H. Heinemann, E. Leipold, Conotoxins of the O-superfamily affecting voltage-gated sodium channels, *Cell. Mol. Life Sci.* 64 (2007) 1329–1340.
- [14] J. Payandeh, T. Scheuer, N. Zheng, W.A. Catterall, The crystal structure of a voltage-gated sodium channel, *Nature* 475 (2011) 353–358.
- [15] W.A. Catterall, A.L. Goldin, S.G. Waxman, International Union of Pharmacology. XLVII. Nomenclature and structure–function relationships of voltage-gated sodium channels, *Pharmacol. Rev.* 57 (2005) 397–409.
- [16] S. Zorn, E. Leipold, A. Hansel, G. Bulaj, B.M. Olivera, H. Terlau, S.H. Heinemann, The μ O-conotoxin MrVIA inhibits voltage-gated sodium channels by associating with domain-3, *FEBS Lett.* 580 (2006) 1360–1364.
- [17] E. Leipold, A. Hansel, B.M. Olivera, H. Terlau, S.H. Heinemann, Molecular interaction of δ -conotoxins with voltage-gated sodium channels, *FEBS Lett.* 579 (2005) 3881–3884.
- [18] J. Barbier, H. Lamthanh, F. Le Gall, P. Favreau, E. Benoit, H. Chen, N. Gilles, N. Ilan, S.H. Heinemann, D. Gordon, A. Menez, J. Molgo, A δ -conotoxin from *Conus ermineus* venom inhibits inactivation in vertebrate neuronal Na⁺ channels but not in skeletal and cardiac muscles, *J. Biol. Chem.* 279 (2004) 4680–4685.
- [19] S. Peigneur, M. Paolini-Bertrand, H. Gaertner, D. Biass, A. Violette, R. Stöcklin, P. Favreau, J. Tytgat, O. Hartley, δ -conotoxins synthesized using an acid-cleavable solubility tag approach reveal key structural determinants for Nav subtype selectivity, *J. Biol. Chem.* 289 (2014) 35341–35350.
- [20] C.I. Schroeder, D.J. Craik, Therapeutic potential of conopeptides, *Future Med. Chem.* 4 (2012) 1243–1255.
- [21] L. Volpon, H. Lamthanh, J. Barbier, N. Gilles, J. Molgo, A. Menez, J.M. Lancelin, NMR solution structures of delta-conotoxin EVIA from *Conus ermineus* that selectively acts on vertebrate neuronal Na⁺ channels, *J. Biol. Chem.* 279 (2004) 21356–21366.
- [22] S. Chierici, M. Jourdan, M. Fiquet, P. Dumy, A case study of 2,2-dimethylthiazolidine as locked cis proline amide bond: synthesis, NMR and molecular modeling studies of a δ -conotoxin EVIA peptide analog, *Org. Biomol. Chem.* 2 (2004) 2437–2441.
- [23] A.A. Miloslavina, E. Leipold, M. Kijas, A. Stark, S.H. Heinemann, D. Imhof, A room temperature ionic liquid as convenient solvent for the oxidative folding of conopeptides, *J. Pept. Sci.* 15 (2009) 72–77.
- [24] P. Heimer, A.A. Tietze, M. Böhm, R. Giernoth, A. Kuchenbuch, A. Stark, E. Leipold, S.H. Heinemann, C. Kandt, D. Imhof, Application of room-temperature aprotic and protic ionic liquids for oxidative folding of cysteine-rich peptides, *Chembiochem* 15 (2014) 2754–2765.
- [25] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins* 59 (2005) 687–696.
- [26] M.S. Cheung, M.L. Maguire, T.J. Stevens, R.W. Broadhurst, DANGLE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure, *J. Magn. Reson.* 202 (2010) 232–233.
- [27] E. Harjes, S. Harjes, S. Wohlgemuth, K.H. Müller, E. Krieger, C. Herrmann, P. Bayer, GTP-Ras disrupts the intramolecular complex of C1 and RA domains of Nore1, *Structure* 14 (2006) 881–888.
- [28] E. Krieger, T. Darden, S.B. Nabuurs, A. Finkelstein, G. Vriend, Making optimal use of empirical energy functions: force-field parameterization in crystal space, *Proteins: Struct., Funct., Bioinf.* 57 (2004) 678–683.
- [29] E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, K. Karplus, Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8, *Proteins: Struct., Funct., Bioinf.* 77 (2009) 114–122.
- [30] E. Krieger, G. Koraimann, G. Vriend, Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field, *Proteins: Struct., Funct., Bioinf.* 47 (2002) 393–402.
- [31] E. Krieger, J.E. Nielsen, C.A. Spronk, G. Vriend, Fast empirical pKa prediction by Ewald summation, *J. Mol. Graph. Model.* 25 (2006) 481–486.
- [32] M.M. Pathak, V. Yarov-Yarovoy, G. Agarwal, B. Roux, P. Barth, S. Kohout, F. Tombola, E.Y. Isacoff, Closing in on the resting state of the shaker K⁺ channel, *Neuron* 56 (2007) 124–140.
- [33] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25 (1997) 3389–3402.
- [34] R.W. Hoof, C. Sander, M. Scharf, G. Vriend, The PDBFINDER database: a summary of PDB, DSSP and HSSP information with added value, *Comput. Appl. Biosci.* 12 (1996) 525–529.
- [35] A.A. Canutescu, A.A. Shelenkov, R.L. Dunbrack Jr., A graph theory algorithm for rapid protein side chain prediction, *Protein Sci.* 12 (2003) 2001–2014.
- [36] E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker, K. Karplus, Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8, *Proteins: Struct., Funct., Bioinf.* 77 (2009) 114–122.
- [37] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593.
- [38] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G.M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J.M. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, *J. Comput. Chem.* 24 (2003) 1999–2012.
- [39] O. Trott, A.J. Olson, Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455–461.
- [40] E. Krieger, G. Vriend, New ways to boost molecular dynamics simulations, *J. Comput. Chem.* 36 (2015) 996–1007.
- [41] J.S. Trimmer, S.S. Cooperman, S.A. Tomiko, J.Y. Zhou, S.M. Crean, M.B. Boyle, R.G. Kallen, Z.H. Sheng, R.L. Barchi, F.J. Sigworth, et al., Primary structure and functional expression of a mammalian skeletal muscle sodium channel, *Neuron* 3 (1989) 33–49.
- [42] H. Chen, D. Gordon, S.H. Heinemann, Modulation of cloned skeletal muscle sodium channels by the scorpion toxins Lqh II, Lqh III, and Lqh α IT, *Pflügers Arch.* 439 (2000) 423–432.
- [43] O. Buczek, B.R. Green, G. Bulaj, Albumin is a redox-active crowding agent that promotes oxidative folding of cysteine-rich peptides, *Pept. Sci.* 88 (2007) 8–19.
- [44] D.J. Craik, N.L. Daly, C. Waine, The cystine knot motif in toxins and implications for drug design, *Toxicol.* 39 (2001) 43–60.
- [45] L.Q. Chen, V. Santarelli, R. Horn, R.G. Kallen, A unique role for the S4 segment of domain 4 in the inactivation of sodium channels, *J. Gen. Physiol.* 108 (1996) 549–556.
- [46] K.J. Kontis, A.L. Goldin, Sodium channel inactivation is altered by substitution of voltage sensor positive charges, *J. Gen. Physiol.* 110 (1997) 403–413.
- [47] W. Stühmer, F. Conti, H. Suzuki, X.D. Wang, M. Noda, N. Yahagi, H. Kubo, S. Numa, Structural parts involved in activation and inactivation of the sodium channel, *Nature* 339 (1989) 597–603.
- [48] F.J. Tejedor, W.A. Catterall, Photoaffinity labeling of the receptor site for α -scorpion toxins on purified and reconstituted sodium channels by a new toxin derivative, *Cell. Mol. Neurobiol.* 10 (2007) 257–265.
- [49] W.J. Thomsen, W.A. Catterall, Localization of the receptor site for α -scorpion toxins by antibody mapping: implications for sodium channel topology, *Proc. Natl. Acad. Sci. U.S.A.* 86 (1989) 10161–10165.
- [50] E. Leipold, A. Borges, S.H. Heinemann, Scorpion beta-toxin interference with Nav channel voltage sensor gives rise to excitatory and depressant modes, *J. Gen. Physiol.* 139 (2012) 305–319.

- [51] [E. Leipold, S. Zorn, H. De Bie, A. Borges, B.M. Olivera, H. Terlau, S.H. Heinemann, Mu O-conotoxins inhibit NaV channels by interfering with their voltage sensors in domain-2, *Biophys. J.* \(2007\) 178A.](#)
- [52] [J. Wang, V. Yarov-Yarovoy, R. Kahn, D. Gordon, M. Gurevitz, T. Scheuer, W.A. Catterall, Mapping the receptor site for \$\alpha\$ -scorpion toxins on a Na⁺ channel voltage sensor, *Proc. Natl. Acad. Sci. U. S. A.* 108 \(2011\) 15426–15431.](#)
- [53] [X. Zhang, W. Ren, P. DeCaen, C. Yan, X. Tao, L. Tang, J. Wang, K. Hasegawa, T. Kumasaka, J. He, D.E. Clapham, N. Yan, Crystal structure of an orthologue of the NaChBac voltage-gated sodium channel, *Nature* 486 \(2012\) 130–134.](#)
- [54] [J. Payandeh, T.M. Gamal El-Din, T. Scheuer, N. Zheng, W.A. Catterall, Crystal structure of a voltage-gated sodium channel in two potentially inactivated states, *Nature* 486 \(2012\) 135–139.](#)
- [55] [Y. Jiang, A. Lee, J. Chen, V. Ruta, M. Cadene, B.T. Chait, R. MacKinnon, X-ray structure of a voltage-dependent K⁺ channel, *Nature* 423 \(2003\) 33–41.](#)
- [56] [J. Wu, Z. Yan, Z. Li, C. Yan, S. Lu, M. Dong, N. Yan, Structure of the voltage-gated calcium channel Cav1.1 complex, *Science* 350 \(2015\) aad2395.](#)
- [57] [K.A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M.A. Wilson, G.A. Petsko, M. Karplus, C.G. Hubner, D. Kern, Intrinsic motions along an enzymatic reaction trajectory, *Nature* 450 \(2007\) 838–844.](#)
- [58] [K.A. Henzler-Wildman, M. Lei, V. Thai, S.J. Kerns, M. Karplus, D. Kern, A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, *Nature* 450 \(2007\) 913–927.](#)
- [59] [J.C. Jones, A.M. Jones, B.R. Temple, H.G. Dohlman, Differences in intradomain and interdomain motion confer distinct activation properties to structurally similar G alpha proteins, *Proc. Natl. Acad. Sci. U. S. A.* 109 \(2012\) 7275–7279.](#)
- [60] [R. Chen, S.H. Chung, Conserved functional surface of antimammalian scorpion \$\beta\$ -toxins, *J. Phys. Chem. B* 116 \(2012\) 4796–4800.](#)
- [61] [R. Chen, S.H. Chung, Binding modes and functional surface of anti-mammalian scorpion \$\alpha\$ -toxins to sodium channels, *Biochemistry* 51 \(2012\) 7775–7782.](#)

Epilogue

In this manuscript the structure-activity relationships of δ -EVIA at different VGSCs were studied in order to understand the mode of action concerning channel specificity. Therefore, the NMR structure has been re-determined and the disulfide connectivity confirmed as the native one by using LC-ESI-MS/MS after partial reduction and derivatization. The NMR structure then has been used in molecular modelling studies to identify the binding site at VGSCs. Homology modelling, MD simulations and docking of the conotoxin was applied to the unaffected channel Na_v1.4, to Na_v1.7, where fast channel inactivation is observed, and to different chimeras of both channels. Considering as well the results from electrophysiological experiments using the same channels and channel chimeras the impact of each domain to the binding of δ -EVIA was identified, and domain I and IV determined to be essential for the specific interaction with the neuronal Na_v1.7.

4.5 Computational study on μ -conotoxin conformations using molecular dynamics and computer visualization methods

Full Paper

Authors

Ajay Abisheck Paul George, Pascal Heimer, Raphael Menges, Veda Thota, Astrid Maaß, Jan Hamaekers, Martin Hofmann-Apitius, Nils Lichtenberg, Kai Lawonn, Diana Imhof

Submitted to ChemPhysChem, September 2017.

Preface

The folding of disulfide-rich peptides seems to be highly sequence-dependent and thus, was investigated with various experimental and theoretical approaches.^{63,109} In this study the folding of a set of μ -conotoxins used as model peptides was in the focus of research. First, the folding was monitored experimentally in buffer solution employing redox-active agents. In general, folding behaviour of these conotoxins is examined starting from the reduced to oxidized peptide with the respective biological activity. Here, MD simulations using the existing NMR structures were conducted to observe the impact of a distinct disulfide bond on the stability of the native conformation compared to partially folded peptide by successive removal of distinct disulfide bonds. Furthermore, a visualistics approach rendering the structure into different layers was utilized for a more global view on the fluctuation of individual residues and their surface proximity throughout the simulation.

Computational study on μ -conotoxin conformations using molecular dynamics and computer visualization methods

Ajay Abisheck Paul George,^[a] Pascal Heimer,^[a] Raphael Menges,^[b] Veda Thota,^[a] Astrid Maaß,^[c] Jan Hamaekers,^[c] Martin Hofmann-Apitius,^[d,e] Nils Lichtenberg,^[b] Kai Lawonn,^[b] Diana Imhof^{*,[a]}

Abstract: The study of protein conformations using molecular dynamics (MD) simulations has been in place for decades. A major contribution to structural stability and native conformation of a protein is made by the primary sequence and disulfide bonds formed during the folding process. Here we studied the μ -conotoxins GIIIA, KIIIA, PIIIA, SIIIA, and SmIIIA as model peptides possessing three disulfide bonds. Their NMR structures were used for MD simulations in a novel approach. The simulation was conducted at a point between the folded native and the unfolded conformation by systematically breaking the disulfide bonds. The main objective was to monitor the cysteines regarding their behavior to support conformational stability and re-folding to the native conformation. An attempt was made to correlate the properties of the primary sequence with preferred folded conformations. In addition, a new visualization approach on a residue's location within a molecule was established that reveals novel insights into residue contribution to structural stability.

1. Introduction

Conotoxins are neuropeptides from the venom of marine cone snails, which interact with a wide range of biological targets (e.g. ion channels, transmembrane receptors, transporters) and hence are of pharmaceutical interest and of great potential as molecular probes to study specific subtypes of ion channels and receptors.^[1]

Conotoxins represent peptides of approximately 10-50 amino acid residues, classified according to their cysteine patterns and disulfide connectivity.^[2,3] The typical CC-C-C-CC pattern defines the framework for conotoxins of the M-superfamily comprised by ψ -, μ -, and κ M-conotoxins.^[2,3] The family of the 27 currently

known μ -conotoxins^[4], selectively binds to the ion channel pore of voltage-gated sodium channels (VGSCs), thus blocking the influx of sodium ions into the cell. Therefore, these peptides also gained interest as useful tools for research studies in electrophysiology.^[3,5] The key determinant for their successful implementation, however, is an in-depth knowledge of the respective structure-activity relationships, which requires extraction of structural and biological data obtained from usually scarce quantities of material. The formation of multiple disulfide bonds in such peptides often hampers easy access by a synthetic approach.^[3,5]

μ -Conotoxins are known to primarily form the linkage C1-C4, C2-C5, and C3-C6, also referred to as the "native fold".^[6-8] The major question here is: Why do different primary amino acid sequences predominantly form a distinct fold out of fifteen possible disulfide isomers?^[6] The molecular principles underlying this observation as well as the folding process remain elusive so far. In addition, prediction of the folding process is hampered due to a lack of adequate data. The subject however, is of a much greater importance than is generally recognized, since an accumulation of cysteines may also occur in distinct regions of larger peptides and proteins as can be exemplified by defensins,^[9] resistin,^[10] Kunitz serine protease inhibitors^[11] and various growth factors.^[10,12,13] In such cases information about the preferred disulfide connectivity based on the amino acid sequence might be important if neither X-ray data nor solution NMR structure analysis is available, which leads to the still unresolved problem of protein folding.

In general, the protein folding problem is discussed to comprise three issues, i.e. (i) the thermodynamic code for a given primary sequence to fold, (ii) the possibility of predicting a protein's native structure from its amino acid sequence, and (iii) the underlying kinetic process depending on the folding pathway.^[14,15] Conotoxins represent a promising tool for the elucidation of folding processes because of their small to medium size and high disulfide bond content representing intermediates between peptides and proteins. Furthermore, they are natural peptides interacting with ion channels and receptors and as such interesting pharmacological tools.^[2,3,5]

Therefore, we studied the μ -conotoxins GIIIA, KIIIA, PIIIA, SIIIA and SmIIIA from a new perspective using molecular dynamics (MD) simulation and a novel visualization tool based on the NMR structures of these compounds. For comparison of experimental and theoretical data, we repeated the oxidation reaction for all μ -conotoxins as described earlier.^[6] Reaction progress (self-folding) was monitored in a conventional buffer system containing redox agents. The individual sequences turned out to be distinguishable regarding reaction product yield and quality similar to what has been reported earlier.^[7] With respect to the MD studies a rather unconventional novel approach was pursued by analyzing the reversed process, i.e. the unfolding of these peptides. More precisely, the behavior and stability of an individual peptide NMR structure was observed upon successive disulfide bond opening. Analyses

[a] A.A. Paul George,* Dr. P. Heimer,* V. Thota, Prof. Dr. D. Imhof
Pharmaceutical Biochemistry and Bioanalytics,
Pharmaceutical Institute, University of Bonn,
An der Immenburg 4, D-53121 Bonn (Germany)
E-mail: dimhof@uni-bonn.de

[b] R. Menges,* N. Lichtenberg, K. Lawonn
Medical Visualization, Institute for Computational Visualistics,
University of Koblenz-Landau, Campus Koblenz,
Universitätsstraße 1, D-56070 Koblenz (Germany)

[c] Dr. A. Maaß, Dr. J. Hamaekers
Department of Virtual Material Design,
Fraunhofer Institute for Algorithms and Scientific Computing,
Schloss Birlinghoven, D-53754 Sankt Augustin (Germany)

[d] Prof. Dr. M. Hofmann-Apitius
Department of Bioinformatics,
Fraunhofer Institute for Algorithms and Scientific Computing,
Schloss Birlinghoven, D-53754 Sankt Augustin (Germany)

[e] Prof. Dr. M. Hofmann-Apitius
Bonn-Aachen International Center for Information Technology,
University of Bonn,
Dahlmannstr. 2, D-53113 Bonn (Germany)

[+]
These authors contributed equally to this work.
Supporting information for this article is given via a link at the end of the document.

conducted on the resulting MD trajectories gave rise to inferences on characteristic factors contributing to conformational stability of the folded conopeptide, e.g. formation of helices. On the other hand, we introduced a novel visualization method based on a geometric approach. The premise of this approach is the extraction of information about a residue's location with respect to the protein's surface, thus rendering the analysis of the MD trajectories more visual and tangible. The results of this work firstly group the five μ -conotoxins based on the stability of their disulfide-deficient versions and also contributes general guidelines before venturing into synthesis.

2. Results and Discussion

2.1. Oxidative self-folding of μ -conotoxins

In order to provide experimental data for comparison, equal amounts of each μ -conotoxin studied herein (Table S1) were used for the oxidation reaction according to a protocol earlier described.^[6] Under the conditions applied μ -SmIIIA and μ -PIIIA oxidation resulted in several peaks of fully oxidized product(s), i.e. identity of the individual fractions is different from the native fold (C1-C4, C2-C5, C3-C6).^[6] In contrast, oxidation of all other peptides, e.g. μ -GIIIA, μ -KIIIA and μ -SIIIA, resulted in one major product as demonstrated earlier (Figure S1).^[7,16,17]

The formation of one main product for μ -GIIIA and μ -SIIIA can be explained by a rapid collapse into the favored "native" connectivity.^[7] For μ -KIIIA such a rapid collapse results also in one main product, however, the connectivity C1-C5, C2-C4, C3-C6 is preferred here, and the "native" fold is only present as a minor fraction.^[18] Small differences in the elution profiles of buffer-oxidized μ -KIIIA (crude product), i.e. product formation, might result from differences in batch size and composition of the oxidation buffer compared to Khoo *et al.*^[18] For instance, a 2:1 ratio of reduced and oxidized glutathione compared to a 1:1 ratio described by Khoo and coworkers^[18] was used in our study. In case of μ -PIIIA and μ -SmIIIA, a different folding mechanism indicative of a slower re-arrangement results in the formation of several isomers.^[7,19] This can be attributed to more diverse noncovalent interactions and electrostatic forces compared to μ -GIIIA and μ -SIIIA. Here, it was suggested that the native isomer accumulates via re-shuffling of disulfide bonds during the folding process and is dependent on the thermodynamic stability of the isomer formed.^[7,19] Although μ -SIIIA and μ -SmIIIA have a high sequence similarity (Table S1), the higher number of basic residues in μ -SmIIIA (six Arg) results in the formation of multiple isomers, while in μ -SIIIA (two Arg and one Lys) only one isomer is preferred.^[7] μ -PIIIA forms multiple isomers for the same reason.^[6] In contrast, the structure of μ -GIIIA tolerates a high number of basic residues and forms only one major product compared to μ -PIIIA and μ -SmIIIA.^[7] Several aspects influencing the folding, such as the number of hydroxyproline residues, loop size between two linked cysteines or amidation of the C-terminus are discussed controversially in the literature without a clear preference indicating the uniqueness of each sequence and respective biological activity.^[7,8,20,21]

2.2. Conformational analysis using molecular dynamics

The solution NMR structures used herein were derived from the *ConoServer* database (Figure 1, Table S1).^[4] Comparison, alignment and structural differences of the μ -conotoxin structures were discussed previously by Yao *et al.* and Tietze *et al.*^[6,22] All μ -conotoxins, except μ -KIIIA, align significantly better in the C-terminal part compared to the N-terminal revealing a conserved structure.^[22] High similarity was shown for the backbone conformations between loop 2 and loop 3 regions of μ -KIIIA, μ -SIIIA, and μ -SmIIIA which have a higher selectivity for blocking Nav1.2 over Nav1.4 channels. On the other hand, μ -GIIIA and μ -PIIIA which prefer Nav1.4 over Nav1.2 superimpose well in the region of the second loop between C2 and C5.^[6] Besides contributing to the structural rigidity, the disulfide bridges cause cysteine residues to form a hydrophobic core, enveloped by other charged and hydrophilic residues (Figure 1). This hydrophobic effect plays a key role in the stability of the native fold.^[21]

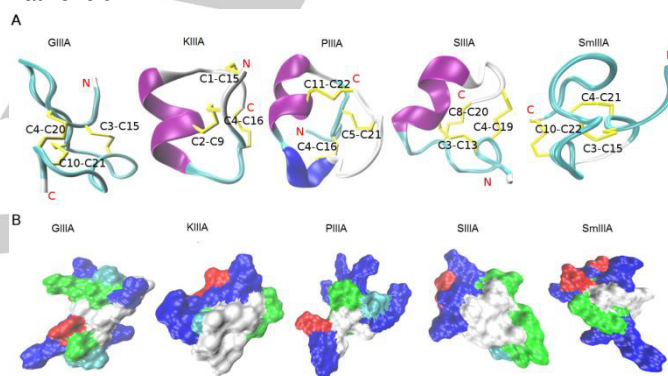


Figure 1. Structure and surface representations of the investigated μ -conotoxins. (A) The NMR structures of the five μ -conotoxins (GIIIA, KIIIA, PIIIA, SIIIA and SmIIIA) used in this study represented as cartoon. Secondary structure elements, α -helix (purple), 3-10 helix (blue), turn (cyan) and coil (white) were generated by STRIDE in VMD.^[23] The cysteine residues forming the disulfide bonds (yellow) are labeled. (B) Molecular surface was generated by SURF in VMD^[23] indicating the hydrophobic (white), basic (blue), acidic (red) and hydrophilic (green) regions.

Regarding the simulation strategy it was decided to open the longest disulfide bridge first (biggest loop) as it would instantly introduce the highest level of flexibility into the peptide backbone. The intention for this opening strategy was to increase conformational entropy of the reduced version. In the case of μ -KIIIA, the shortest disulfide bridge was opened first. This would serve as a means to gauge the effect of loop size in retaining a stable structure. The loops in between the three disulfide bonds differ in size (Table S2) and thus, need to be considered in the analysis process.

It was observed during the simulations at room temperature that all natively folded structures showed good conformational stability based on the evaluation of root mean square deviation (RMSD) compared to NMR ensemble of each peptide with μ -KIIIA possessing the lowest RMSD of 1.1 Å and μ -SmIIIA with the highest of 2.7 Å (Figure S2, Table S3). A test simulation of μ -

PIIIA was done for 1 μ s, and we considered that using 100 ns of the trajectory was a good representation of the simulation (Figure S3). RMSD of the backbone, root mean square of fluctuation (RMSF) of all atoms of each residue and the radius of gyration (Rg) of the whole protein were computed from each simulation for all of the five peptides (Figure S2, S4).

The native structure of μ -GIIIA was observed to be held together by a combination of forces. Apart from the evident hydrophobic core formed by the cysteines, the presence of aspartic acid in its sequence gave rise to salt bridge formation observed between R1 and D12. As the C2-C5 disulfide bond was removed an increase in distance between the now unbound cysteines was observed. However, this did not affect the stability of the backbone proved by only a 0.3 Å increase in backbone RMSD between the native structure and the one with the C2-C5 disulfide removed, but it could be observed that its RMSD progression is inconsistent in the simulation especially between 30 ns and 50 ns. The peptide had excessive flexibility and adopted close to completely unfolded conformations during different intervals in the simulation (Figure S5). Importantly, the residues significant for bioactivity experienced none to very minimal increase in fluctuations (Table S4). Upon subsequent removal of the C3-C6 disulfide bridge, an obvious stretch in the overall shape of the peptide was observed. Interestingly, the cysteines forming the C3-C6 bridge moved much closer to each other than the C2-C5 cysteine residues in the two disulfide bridges opened structure. This is shown by the decrease in RMSD of the C2-C5 and C3-C6 removed peptide between 60–70 ns (Figure S5). Rg of μ -GIIIA followed an almost identical pattern of progression to the RMSD, peaking between 40 and 55 ns before falling back towards its initial values, representing an unfolding-(re)folding event. The fluctuation of the residues K8, K11, R13, K16 and R19, which were reported to be responsible for binding,^[3] did not show significant change in fluctuation compared to the structure with all disulfide bridges intact (Figure S2). With two disulfide bridges removed, R19 and K16 showed significant increase in fluctuations. It has been shown that K16 has a low priority for biological activity. Moreover, the exchange of this residue increased binding affinity compared to the native toxin.^[24] However, usually alanine residues confer higher hydrophobicity into the peptide backbone. Replacement of cysteine by serine would probably be a better choice regarding retaining the bioactivity. It is safe to state that if the C2-C5 disulfide bond alone was removed the structure adopted a conformation still representing reasonable stability.

μ -KIIIA has the lowest RMSD of 1.1 Å among the five peptides in their natively folded form. Removal of the C2-C4 disulfide bond thereby altering the C4 residue results in the loss of its helix and increases the RMSD by 0.7 Å. The remaining C1-C5 and C3-C6 disulfide bonds, however, were sufficient to retain backbone stability of μ -KIIIA. The progression of RMSD and Rg for μ -KIIIA followed the same scheme as observed for μ -GIIIA, i.e. with the structure possessing two reduced disulfides revealing the largest variation in conformational flexibility.

Although the average RMSD of native μ -PIIIA was a decent 2.3 Å over the course of the simulation, the backbone was constantly subjected to changes as can be seen from the RMSD

plot (Figure S2). The removal of the C2-C5 disulfide bond resulted in a lower fluctuating RMSD progression although it came at the expense of 1.5 Å increase in RMSD. The structure of μ -PIIIA with both the C2-C5 and C3-C6 disulfide bridges removed showed the largest extent of structural variation among the five conopeptides with an average RMSD of 4.9 Å (Figure S6). The peptide did not tend to refold within 100 ns simulation time. With one disulfide removed, none of its functionally significant residues showed pronounced increase in fluctuations, the highest of which was a 1.3 Å increase for R14 compared to the natively folded peptide. In contrast, almost all residues of μ -SIIIA displayed marginally higher RMSF values for the one and two bond-removed structures in comparison to the native structure containing all three disulfide bonds.

We focused on μ -GIIIA that forms a single, and μ -SmIIIA that forms multiple oxidation products during synthesis to illustrate this phenomenon described above (Figure S5). The overall RMSD between the native connectivity and the structure with one disulfide bond removed was observed to be the lowest in μ -GIIIA (1.8 Å) and highest in μ -SmIIIA (2.7 Å) among the five peptides. Figure S4 shows average conformations for all three simulations of μ -GIIIA and μ -SmIIIA compared with their corresponding RMSD plots. An interesting observation from both peptides with opened disulfide bond was the formation of new secondary structure elements that were not present in the native state. In μ -SmIIIA with one opened disulfide bond, an α -helix was formed between residues R13 and H18 (Figure 2).

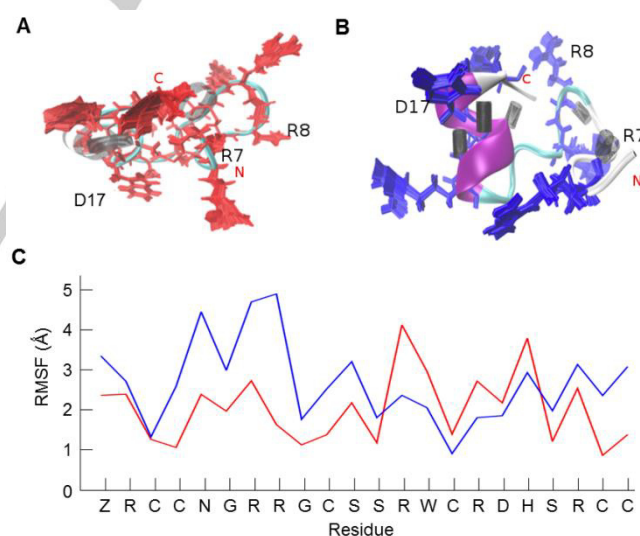


Figure 2. Comparison of μ -SmIIIA native fold and the structure with one disulfide bond opened. (A) The structure of the native peptide (completely oxidized, three disulfide bonds) with 100 conformations of its basic residues (red) superimposed with three residues marked as important for binding activity.^[3] The distribution of hydrogen bonds shows a sparse black area which indicates that the region surrounding it is relatively flexible. (B) The structure of the peptide containing one opened disulfide bond (C2-C5) showing (blue) its well-formed α -helix and the dense well-ordered hydrogen bonding illustrated as black cylinders. Higher rigidity, inducing improved structural stability of the peptide B with the C2-C5 disulfide bond removed in comparison to the fully oxidized peptide with all three disulfide bonds is apparent from the reference RMSF plot (C). Despite the rigidity of the backbone the orientations of the basic residues differs from the native structure.

The stability of this helix through the entire course of the simulation can be accounted for by a combination of hydrogen bond formation and the presence of the bonded C15 in the center of the helix. In comparison with the native fold, the structure with the C2-C5 bond opened appeared well ordered. The distribution of the hydrogen bonds around the helix can be seen in Figure 2.

The μ -SmIIIA structure with two disulfide bonds reduced formed a less stable 3_{10} helix between residues R16 and H18 increasing the flexibility of the conformation. However, functionality of a conotoxin is dependent on a stable backbone structure coupled with the favorable orientation of basic side chain residues for binding to their target and not solely on the flexibility of a distinct part of the peptide.^[3] In μ -SmIIIA, though the 3_{10} helix formation reduced Rg of the peptide, the overall peptide structure drifted significantly from that of the native state, and the orientations adopted by the side chains of its basic residues varied largely when compared to either the native or the structure with one disulfide bond opened. Even with a single disulfide removed the functionally significant R7 showed a moderate increase in fluctuations (Table S4).

Herewith it is demonstrated how the individual disulfide bridges and the residues that occur within the loops affect the conformational stability on the basis of observed fluctuations in the backbone and the side chain residues. In the following section a shift in perspective in the extraction of useful information from MD trajectories is introduced. Here we focus on global atomic movements within the molecule and how these contribute to observable surface level changes compared to local molecular fluctuations.

2.3. Computation of residue surface proximity

This study was performed with the aim to find a measure for each amino acid that gives us an impression of how close that particular residue is to the molecule's surface. This measure was called "Residue Surface Proximity". First, we computed an atom-layer extraction based on the solvent-accessible surface (SAS)^[25] to assign each atom to a specific surface layer (Figure 3, S7). The idea of layers was inspired by the work of Karampudi *et al.*^[26] This method is able to assign atoms and their affiliated residues to a variation of layers. The different layers consist of the outermost surface layer, the intermediate layers, and a core layer. In general, an amino acid is assigned to the innermost layer that its atoms extend into. If an amino acid is found at the boundary of two layers, it is assigned to the inner layer, towards the core. Similar considerations were utilized in our approach (Figure 3). We adapted this approach according to the objectives, i.e. to make it feasible for the application to small proteins or peptides, such as μ -conotoxins. Our goal was to use the layer-membership of a residue as a hint for its surface proximity. The change of the layer-membership was used as a measure of stability, with respect to the surface proximity, throughout an MD simulation. In addition, a GPU implementation of an algorithm described by Zhang *et al.*^[27] was used to extract

the surface atoms. This algorithm identified all atoms that are part of the SAS of a given input molecule. Thus, we found the outer layer of the input data based on a selected probe radius (see also Methods) with respect to the SAS definition. All remaining atoms were classified as internal. The algorithm was then repeated using the atoms previously classified as internal layer as input, which yielded the next layer of atoms. The algorithm stopped once all atoms were part of a particular layer, e.g. after N iterations (Figure 3).

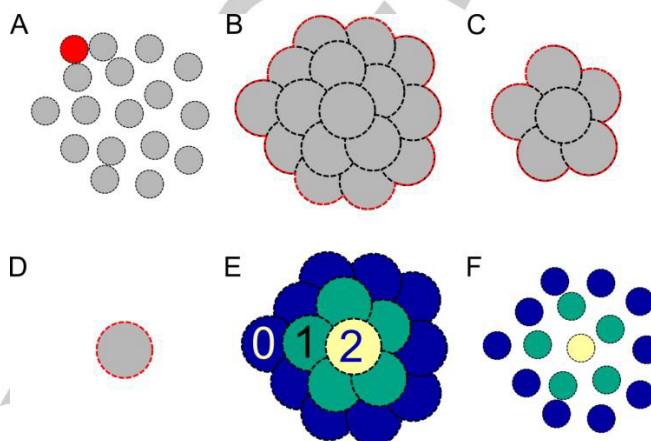


Figure 3. Assembly of atoms in 2D space with their Van der Waals radius (gray) and a SAS probe (red, A). (B) Extending the atom radii by the probe radius yields the extended hull and the SAS (marked red). The outer atoms form the first SAS layer. (C) The outer atoms of the first layer are removed and the next layer of SAS is formed. (D) This process was repeated until the last layer was reached. (E) Each atom was assigned a value denoted as L_A that was based on the iteration in which the atom has been assigned to the SAS. (F) Finally, an assignment of layers to each atom was obtained based on the definition of the SAS.

The trajectories of the 100 ns MD were used as input to our layer extraction software. For the SAS definition a probe radius of 2.8 Å was used. This radius was chosen twice as high as the commonly used value of 1.4 Å,^[27,28] which approximates the size of a solvent water molecule. This decision is justified by layers to represent the surface proximity of each atom rather than the standard SAS. By increasing the probe radius this algorithm produces more layers, which gives a more fine-grained classification. Also, the spatial difference between two adjacent layers is lower as the total number of layers increases. For 1.4 Å, μ -GIIIA (Figure 4A) and μ -SmIIIA (Figure 5A) mostly contained three layers. In case of 2.8 Å, the number of layers increased to 4 to 5. Furthermore, even higher probe radii were investigated up to values of 6 Å, but the layer count did not increase much in these cases. Based on that finding, 2.8 Å was considered for the envisaged computation with an optimum ratio between layer count and SAS representation. The computations of the surface proximity for each amino acid were carried out by our software.

As demonstrated for μ -GIIIA, where the first disulfide bond (C3-C15) is forming the innermost layer over the whole simulation time (Figure 4), only few layer changes indicate a rigid structure and low flexibility of residues. However, for μ -KIIIA the residues are stable within the respective layer (Figure S8). This is due to the short sequence resulting in a more compressed and rigid structure compared to the other peptides. The visualization of the different SAS layers for μ -PIIIA (Figure S8) shows C4 moving to the surface after ~ 10 ns of simulation time and follows a series of shifts from the surface to the core during the simulation. Furthermore, it can be observed that the linked C16 has an almost similar layer-membership pattern to that of C4.

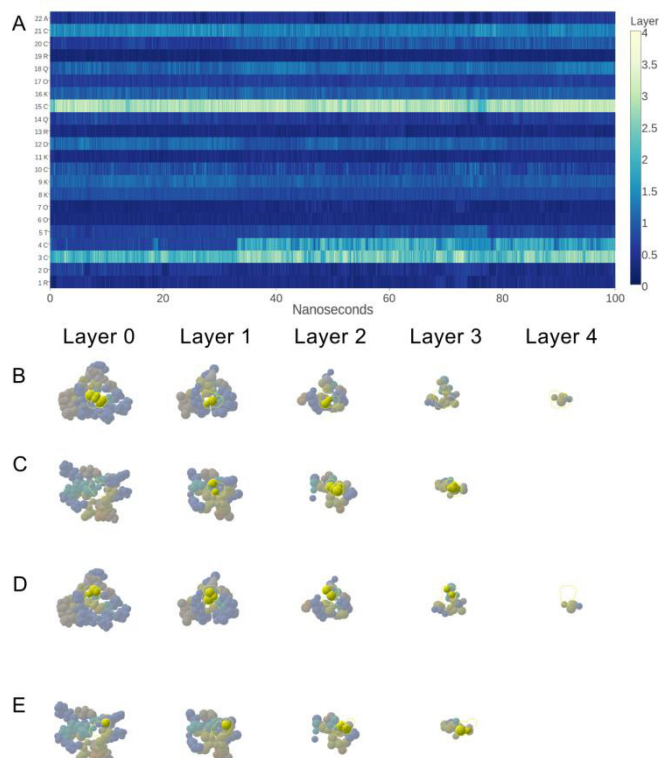


Figure 4. Visualization of the SAS layers at different time steps. (A) represents the layer heatmap for μ -GIIIA. Each column shows the atoms of one of the layers. (B) is the residue C10 captured at 10 ns. (C) is residue C21 captured at 90 ns. (D) is residue C21 captured at 10 ns. (E) is residue C21 captured at 90 ns. This figure proposes that C10 emerges from the center of the molecule towards the surface after ~ 10 ns. It can be observed that both residues (C10 and C21) move to the core at the end of the simulation, which can also be seen directly from the heatmap plot.

Although for μ -SIIIA layer changes can also be observed for the second (C4-C19) and third (C8-C20) disulfide bond, in contrast to μ -PIIIA the average is lower resulting in a more rigid structure again. Finally, μ -SmIIIA represents a more flexible structure because the first disulfide bridge (C3-C15) changes surface availability throughout the simulation (Figure 5). Here the innermost layer is formed by this cysteine bridge, because the two remaining ones (C4-C21, C10-C22) are connected to the cysteines at the end positions of the peptide and are therefore generally surface available. The cysteine residues from the

disulfide bridge C10-C22 are able to change their layer independently. While C22 stays in an outer layer, the connected C10 is only in an outer layer for a short period of time from 40-48 ns before going back to an inner layer (Figure 5).

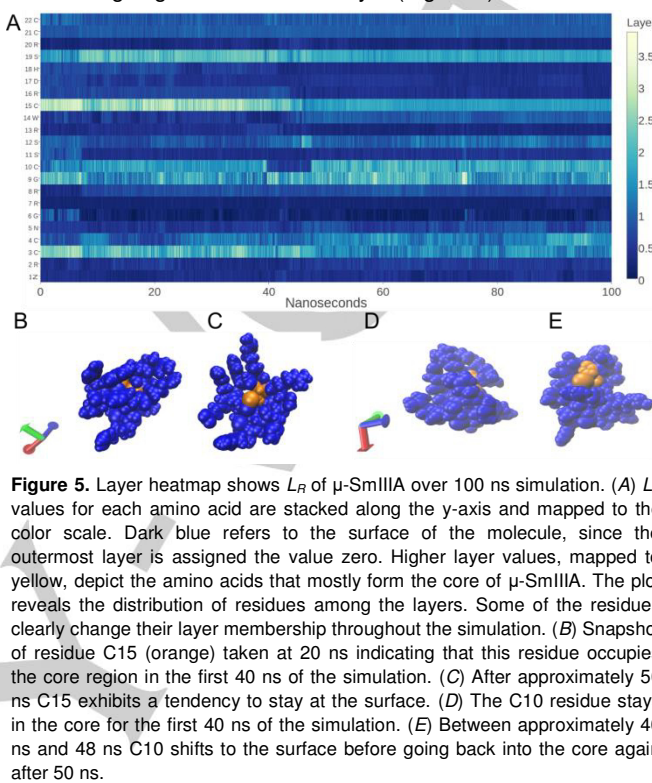


Figure 5. Layer heatmap shows L_R of μ -SmIIIA over 100 ns simulation. (A) L_R values for each amino acid are stacked along the y-axis and mapped to the color scale. Dark blue refers to the surface of the molecule, since the outermost layer is assigned the value zero. Higher layer values, mapped to yellow, depict the amino acids that mostly form the core of μ -SmIIIA. The plot reveals the distribution of residues among the layers. Some of the residues clearly change their layer membership throughout the simulation. (B) Snapshot of residue C15 (orange) taken at 20 ns indicating that this residue occupies the core region in the first 40 ns of the simulation. (C) After approximately 50 ns C15 exhibits a tendency to stay at the surface. (D) The C10 residue stays in the core for the first 40 ns of the simulation. (E) Between approximately 40 ns and 48 ns C10 shifts to the surface before going back into the core again after 50 ns.

Interestingly, also non-cysteine residues such as serine (S19 for μ -SmIIIA, S10 for μ -PIIIA and μ -SIIIA) are able to occupy the core of the peptide. On the other hand, large basic residues (R) remain on the surface throughout the simulation and are thus always available for interaction with the ion channel. The obtained data represents a novel means of relating the surface availability of residues and stability of conotoxins.

3. Conclusions

While protein folding is considered frequently, the folding of smaller disulfide-rich peptides and oligopeptides is less well understood due to a much higher degree of flexibility and often a lower extent of structure-forming elements.^[30] MD simulations have proved to assist experimental work in understanding and predicting the folding process.^[31] Previous studies using the five μ -conotoxins investigated herein gave insights into the folding and binding modes adopted by these peptides.^[6,18] Our work, however, aimed at determining how a particular disulfide bond contributes to the stability of the peptide. Consequently, this approach reviews the validity of the logic that the removal of a disulfide bridge, i.e. herein C2-C5, represents a reduction of the backbone stability when considering RMSD and RMSF values. With the C2-C5 disulfide bond removed, only μ -SmIIIA revealed a noticeable increase in the average fluctuations of its

functionally significant residues. In addition, the fact that secondary structural elements such as α -helices were formed in some peptides containing only two disulfide bonds suggests that in distinct cases (e.g. μ -GIIIA, μ -SmIIIA) greater structural stability of the backbone may be achieved if one disulfide bridge is removed. This helix-induced stability might strengthen the backbone and reduce the extent of overall fluctuations of the basic residues responsible for binding activity.^[3] It can also occur that the nature of a given force field might influence conformation, e.g. induction of helix formation.^[32] Therefore, the cause of retained or gained backbone stability must be always considered. The biological activity and selectivity of disulfide-deficient mutants might differ from the native conformation as shown for μ -GIIIA recently.^[33] 15 different disulfide isomers are possible for a peptide containing six cysteines and still three different isomers (ribbon, bead, and globular) might occur in case of four cysteines.^[34] It has, however, not been mentioned in the report by Han *et al.*, which isomer of the μ -GIIIA analogs has been tested because structural characterization of the respective products was not performed.^[33] Apart from reports on μ -GIIIA regarding disulfide-deficient variants, another study by Khoo *et al.* provides insight into the removal of disulfide bridge C1-C9 in μ -KIIIA resulting in only a minimal change in biological activity against Na_v1.2 and Na_v1.4.^[35] In contrast, there are no experimental data for disulfide-deficient species of μ -conotoxins PIIIA, SIIIA and SmIIIA available so far.

With respect to drug design and synthesis the simplification to two disulfide bonds would be a clear benefit for disulfide-rich peptides and proteins. In this respect, we can conclude from our MD simulations that two disulfide bridges could be sufficient to maintain a stable backbone for the majority of the μ -conotoxins studied. However, it is important that the deficient structure is sufficiently supported by at least one pair of cross-linked disulfide bridges that span to almost either ends of the sequence. From the results obtained a rank order of the five peptides can be provided: μ -GIIIA and μ -KIIIA fall in the highly favorable category, μ -SIIIA falls in the moderately favorable, and μ -PIIA and μ -SmIIIA fall into the least favorable group. We also conclude that the C3-C6 disulfide bridge plays the greatest role in retaining backbone stability.

As expected, removal of two disulfide bridges led to increased backbone flexibility, the formation of a series of intermediate conformations and a less stable peptide. Increased fluctuations of basic side chain residues responsible for interactions with the sodium ion channels are in this case unlikely to stay in favorable orientations for binding. Furthermore, the formation of different disulfide isomers for μ -KIIIA, μ -PIIA and μ -SmIIIA in the experimental self-folding approach indicates a difference in their folding behavior which cannot be explained unambiguously by simulations and disulfide removal.

The second approach, the residue surface proximity, adds a new perspective on residue movements throughout an MD simulation and their influence on structural stability within a peptide. Our findings revealed the importance of considering the effects of global movements within the peptide that might not be apparent as fast thermal vibrations. It is decoupled of absolute spatial movements of the peptide because it only relies on the

layered extraction of the SAS. A residue that changes its layer will very likely affect the layer-membership of other residues and reflects how different amino acids compete with each other. This represents an alternative method in addition to the interpretation of common magnitudes like RMSD, RMSF or Rg, which rely on euclidean space and consider groups of atoms individually. The work by Das *et al.* revealed more information about the counterplay of different residues,^[36] since direction and amplitude of residue movements was derived from a principal component analysis (PCA). In our approach we also received directional and amplitude information. We deduce that the residue surface proximity is an intuitive perspective on MD simulations extracting complex trajectory data more easily. An attempt was made to draw a correlation between the results from the MD simulations and the visualization approach. The correlation between the layer delta ΔL_R and RMSF values revealed interesting features comparing the individual residue flexibility and the invariability of their layer-membership. This can be exemplified with the fact that the majority of amino acids displaying higher flexibility are constantly found on the surface without significantly changing their position, e.g. residues K11, R13 and R19 in μ -GIIIA, K7, W8, R10, R14, in μ -KIIIA, K9, R14, O18, R20 in μ -PIIA, K11, W12, R18 in μ -SIIIA, and R7, S11, R13, W14, R16, D17, H18, and R20 in μ -SmIIIA, ignoring the terminal amino acids. It is not surprising that these residues correlate with the ones important for biological activity. Interestingly, mainly cysteines and serine change their layer-membership more frequently in the peptides. A reason for this might be that the aforementioned basic residues were held in distinct geometric positional range and the cysteine pattern holds the structure together and move to maintain the bioactive conformation. This approach might certainly serve as a tool to study molecular motion and surface properties from a perspective that is not achieved by the usual methods of MD trajectory analysis. The specific focus to global atomic movements within the molecule is a strong visual cue in determining components that contribute to overall stability.

Experimental Section

Chemical synthesis and purification of μ -conotoxins. Peptides were produced by automated solid phase peptide synthesis using a standard Fmoc (*N*-(9-fluorenyl) methoxycarbonyl)-protocol and an EPS 221 peptide synthesizer (Intavis Bioanalytical Instruments AG, Cologne, Germany) as described earlier,^[6] and purified by preparative reversed phase HPLC (Shimadzu LC-8A system, Duisburg, Germany). The gradient used was 0 - 50% eluent B in 120 min with 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile/water (9:1) (eluent B) on a C18 column (Knauer Eurosphere 100, Berlin, Germany) with the dimensions 50 mm x 300 mm (5 mm particle size, 100 Å pore size). Reduced and oxidized peptides were analyzed on a LC-ESI microTOF-Q III mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) coupled with a Dionex Ultimate 3000 (Thermo Scientific, Dreieich, Germany) equipped with a EC100/2 Nucleoshell RP18 Gravity 2.7 μ m column (Macherey-Nagel, Düren, Germany). Analysis of the MS data was performed using Bruker Compass Data Analysis 4.1. LC conditions used were as follows: Eluent A was water with 0.1% acetic acid, while eluent B was acetonitrile

containing 0.1% acetic acid. A gradient of 0% – 60% of eluent B in 12 min was used and detection was at 220 nm.

Oxidation of reduced μ -conotoxin precursors. Oxidative folding of the linear μ -conotoxins GIIIA, KIIIA, PIIIA, SIIIA and SmIIIA in a buffer system containing redox agents was performed as described earlier.^[6] 1 mg of each μ -conotoxin was subjected to oxidation and fractions of the reaction mixture were monitored over time by RP HPLC using a Shimadzu LC-10AT system (Duisburg, Germany) equipped with a C18 column (Vydac 218TP54, Worms, Germany, 4.6 mm x 25 mm, 5 mm particle size, 300 Å pore size) and the gradient 0 - 60% eluent B in 60 min with 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile (eluent B). Reaction control was performed over a time period of 24 h, oxidation was stopped by adding 1% TFA in water. Monitoring revealed that the oxidation reactions were completed within the first 60 min of reaction time. Fractions were collected for each peptide and subjected to LC-ESI mass spectrometry for confirmation of the molar mass corresponding to the oxidized products.

Molecular dynamics (MD) simulations. Disulfide bonds were systematically removed to yield a partially folded, conformation as the starting structure for the simulation. Four out of five μ -conotoxins in this study possess the native disulfide connectivity (C1-C4, C2-C5, C3-C6), whereas μ -KIIIA adopts a C1-C5, C2-C4, C3-C6 as the stable conformation.^[18] The following structures were used herein: PDB ID 1TCG (μ -GIIIA),^[16] PDB ID 2LXG (μ -KIIIA),^[18] S00159 (μ -PIIA),^[6] BMRB 20025 (μ -SIIIA),^[22] and PDB ID 1Q2J (μ -SmIIIA).^[37] Each peptide was subjected to three main production runs and on each run, pre-processing and equilibration was performed independently. First, the conformation with all three disulfide bonds was considered as the control simulation. In the second simulation, the link between C2-C5 (C2-C4 for μ -KIIIA) was removed leaving the two other cysteine bonds intact. In the third simulation, both the C2-C5 (C2-C4 for μ -KIIIA) as well as the C3-C6 disulfide bonds were removed. All peptides underwent energy minimization and equilibration prior to being used as the starting structure of the simulation, as explained in detail below.

GROMACS 5.0.5^[38] was used for all the MD simulations in this study. An individual peptide was placed in the center of a cubic box of size 1 x 1 x 1 nm. TIP3P^[39] water model was used as the solvent to fill the box. Appropriate amounts of Cl⁻ ions were added to balance the positive charge of the μ -conotoxins. Simulations were run using the AMBER99SB-ILDN^[40] force field, which was chosen based on its better agreement with the NMR data and an accurate modeling of helical proteins in comparative studies.^[43,44] In the process of preparing the peptide for the production MD simulation, energy minimization simulations were carried out in three steps. The first step steepest descents minimization with the whole protein held immobile in the solvent, followed by steepest descents with only the backbone frozen. The final step was 5000 steps of steepest descents with all restraints on the peptide removed. This 3-step minimization process with constraints enabled the solvent to soak into all parts of the peptide uniformly resulting in a well-minimized structure for the production MD. A thermal equilibration at 300 K using the modified Berendsen^[43] thermostat and a constant pressure equilibration using the Parrinello-Rahman barostat^[45,46] at 1 atm were carried out for 10ns each, prior to production MD. During both the temperature and pressure ensemble simulations, restraints on all bonds were applied using the LINCS^[46] algorithm. The production MD was done for 100 ns with a 2 fs time step and data written to the logs and trajectory at every 5 ps. Periodic boundary conditions were applied to the system. Long range electrostatics were accounted by the particle mesh Ewald method.^[47,48] 20000 frames were written to the trajectory during each simulation. The effect of periodic boundary conditions was adjusted by suppression of center of mass movement

from the trajectory prior to analysis. Visualizations of conformations for the analysis and creation of images was performed using VMD.^[23]

Molecular surface visualization of global residue mobility. An in-house software prototype was implemented to conduct the layer and residue surface proximity analysis. The prototype includes a GPU implementation of the surface atom extraction algorithm proposed by Zhang.^[27] The algorithm was applied iteratively to compute several layers of atoms, which was inspired by the work of Karampudi and co-workers.^[26] With each layer comes an integer value L_A per atom. L_A was determined by the iteration in which an atom was assigned to a layer. Then, $L_A = 0$ determined the outermost layer and the innermost layer was denoted by atoms with $L_A = N-1$, where N is the number of iterations. For each residue we computed L_R as the mean value of all of its atoms' L_A values. In this way we received a score for each residue that described its proximity to the SAS or the molecule's core, respectively. L_R is now defined as the Residue Surface Proximity. In order to incorporate the dynamics of a MD simulation, we further determined L_R for each residue in each time step x of a simulation and compute $\Delta L_R(x) = |L_R(x) - L_R(x-1)|$. $\Delta L_R(x)$ measures how much an amino acid's surface proximity changes at time x with ΔL_R representing the average change of surface proximity of a residue over the time span of a simulation. ΔL_R can be interpreted as a hint for how stable a residue is in its surface proximity throughout a folding process or how much it competes with other residues. L_R is indicative of how much a residue contributes to the surface or core of a molecule. Our software prototype can be used to visually analyze a residue's surface proximity in further detail.

Acknowledgements

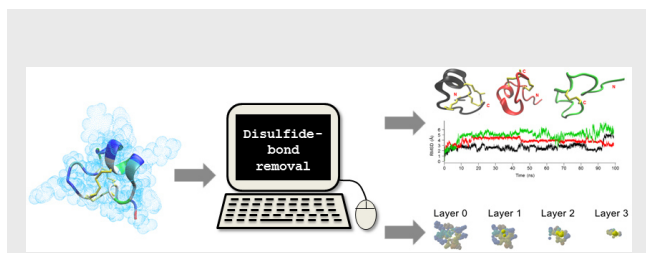
Authors like to thank Alesia A. Tietze and Daniel Tietze (TU Darmstadt) for useful scientific discussions. Financial support by the University of Bonn (to D.I.) and the University of Koblenz-Landau (to K.L.) is gratefully acknowledged.

Keywords: μ -conotoxin • conformational fold • molecular dynamics • computational visualistics

- [1] F. M. Ashcroft, *Nature* **2006**, *440*, 440–447.
- [2] L. J. Cruz, W. R. Gray, D. Yoshikami, B. M. Olivera, *J. Toxicol. Toxin Rev.* **1985**, *4*, 107–132.
- [3] K. B. Akondí, M. Muttenthaler, Q. Kaas, D. J. Craik, R. J. Lewis, P. F. Alewood, *Chem. Rev.* **2014**, *114*, 5815–5847.
- [4] Q. Kaas, R. Yu, A. H. Jin, S. Dutertre, D. J. Craik, *Nucleic Acids Res.* **2012**, *40*, 325–330.
- [5] M. Góngora-Benítez, J. Tulla-Puche, F. Albericio, *Chem. Rev.* **2014**, *114*, 901–926.
- [6] A. A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Kühl, A. Mischo, G. Buntkowsky, M. Görlach, S. H. Heinemann, D. Imhof, *Angew. Chem. Int. Ed.* **2012**, *51*, 4058–4061.
- [7] E. Fuller, B. R. Green, P. Catlin, O. Buczek, J. S. Nielsen, B. M. Olivera, G. Bulaj, *Febs J.* **2005**, *272*, 1727–1738.
- [8] G. Bulaj, B. M. Olivera, *Antioxid. Redox Signal.* **2008**, *10*, 141–155.
- [9] A. Szyk, Z. Wu, K. Tucker, D. Yang, W. Lu, J. Lubkowski, *Protein Sci.* **2006**, *15*, 2749–2760.
- [10] M. S. Lyons, B. Bell, D. Stainier, K. G. Peters, *Dev. Dyn.* **1998**, *212*, 133–140.
- [11] S. Ranasinghe, D. P. McManus, *Dev. Comp. Immunol.* **2013**, *39*, 219–227.
- [12] K. J. Barnham, A. M. Torres, D. Alewood, P. F. Alewood, T. Domagala, E. C. Nice, R. S. Norton, *Protein Sci.* **1998**, *7*, 1738–49.
- [13] H. W. Christinger, G. Fuh, A. M. de Vos, C. Wiesmann, *J. Biol. Chem.* **2004**, *279*, 10382–10388.
- [14] K. A. Dill, S. B. Ozkan, M. S. Shell, T. R. Weikl, *Annu. Rev. Biophys.* **2008**, *100*, 289–316.
- [15] G. D. Rose, P. J. Fleming, J. R. Banavar, A. Maritan, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16623–16633.

- [16] K. Wakamatsu, D. Kohda, H. Hatanaka, J. M. Lancelin, Y. Ishida, M. Oya, H. Nakamura, F. Inagaki, K. Sato, *Biochemistry* **1992**, *31*, 12577–12584.
- [17] D. Patel, S. Mahdavi, S. Kuyucak, *Biochemistry* **2016**, *55*, 1929–1938.
- [18] K. K. Khoo, K. Gupta, B. R. Green, M. M. Zhang, M. Watkins, B. M. Olivera, P. Balaram, D. Yoshikami, G. Bulaj, R. S. Norton, *Biochemistry* **2012**, *51*, 9826–9835.
- [19] A. Walewska, J. J. Skalicky, D. R. Davis, M. M. Zhang, E. Lopez-Vera, M. Watkins, T. S. Han, D. Yoshikami, B. M. Olivera, G. Bulaj, *J. Am. Chem. Soc.* **2008**, *130*, 14280–14286.
- [20] E. Lopez-Vera, A. Walewska, J. J. Skalicky, B. M. Olivera, G. Bulaj, *Biochemistry* **2008**, *47*, 1741–1751.
- [21] T. S. Kang, R. M. Kini, *Cell. Mol. Life Sci.* **2009**, *66*, 2341–2361.
- [22] S. Yao, M. M. Zhang, D. Yoshikami, L. Azam, B. M. Olivera, G. Bulaj, R. S. Norton, *Biochemistry* **2008**, *47*, 10940–10949.
- [23] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.
- [24] G. Choudhary, M. P. Aliste, D. P. Tieleman, R. J. French, S. C. Dudley Jr., *Channels* **2007**, *1*, 344–352.
- [25] B. Lee, F. M. Richards, *J. Mol. Biol.* **1971**, *55*, 379–400.
- [26] N. B. R. Karampudi, R. P. Bahadur, *Sci. Rep.* **2015**, *5*, 16141.
- [27] J. Zhang, Z. Shi, *Comput. Graph.* **2014**, *38*, 291–299.
- [28] N. Lindow, D. Baum, H. C. Hege, *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 2486–2495.
- [29] B. Kozlikova, M. Krone, N. Lindow, M. Falk, M. Baaden, D. Baum, I. Viola, J. Parulek, H.-C. Hege, *Eurographics Conf. Vis.* **2015**, 1–21.
- [30] J. Y. Chang, *Biochemistry* **2011**, *50*, 3414–3431.
- [31] M. Karplus, J. Kuriyan, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6679–6685.
- [32] E. A. Cino, W. Choy, M. Karttunen, *J. Chem. Theory Comput.* **2012**, *8*, 2725–2740.
- [33] P. Han, K. Wang, X. Dai, Y. Cao, S. Liu, H. Jiang, C. Fan, W. Wu, J. Chen, *Mar. Drugs* **2016**, *14*, 1–9.
- [34] J. L. Dutton, P. S. Bansal, R. C. Hogg, D. J. Adams, P. F. Alewood, D. J. Craik, *J. Biol. Chem.* **2002**, *277*, 48849–48857.
- [35] K. K. Khoo, Z. P. Feng, B. J. Smith, M. M. Zhang, D. Yoshikami, B. M. Olivera, G. Bulaj, R. S. Norton, *Biochemistry* **2009**, *48*, 1210–1219.
- [36] A. Das, C. Mukhopadhyay, *J. Chem. Phys.* **2007**, *127*, 165103.
- [37] D. W. Keizer, P. J. West, E. F. Lee, D. Yoshikami, B. M. Olivera, G. Bulaj, R. S. Norton, *J. Biol. Chem.* **2003**, *278*, 46805–46813.
- [38] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19–25.
- [39] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [40] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins Struct. Funct. Bioinf.* **2010**, *78*, 1950–1958.
- [41] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, D. E. Shaw, *PLoS One* **2012**, *7*, 1–6.
- [42] A.-P. Serafeim, G. Salamanos, K. K. Patapati, N. M. Glykos, *J. Chem. Inf. Model.* **2016**, *56*, 2035–2041.
- [43] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, a DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [44] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [45] S. Nosé, M. L. Klein, *Mol Phys* **1983**, *50*, 1055–1076.
- [46] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [47] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J Chem Phys* **1995**, *103*, 8577–8593.
- [48] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.

ARTICLE



Disulfide-bond removal of μ -conotoxin: A combined approach of molecular dynamics simulation and a novel visualistics method based on SAS reveal insights in the stability of distinct disulfide bonds and the positioning of basic amino acids responsible for bioactivity.

A.A. Paul George, P. Heimer, R. Menges, V. Thota, A. Maaß, J. Hamaekers, M. Hofmann-Apitius, N. Lichtenberg, K. Lawonn, D. Imhof*

1 – X.

Computational study on μ -conotoxin conformations using molecular dynamics and computer visualization methods

Epilogue

In this report MD simulations of the μ -conotoxins GIIIA, KIIIA, PIIIA, SIIIIA and SmIIIA were performed to expose the structural stability of the folded and unfolded species with one and two disulfide bonds removed. As a result, the structures were assigned to classes according to their structural stability in disulfide-deficient variations. These were characterized with magnitudes such as RMSD, RMSF and radius of gyration and on the other hand with a layer approach revealing the positions of individual residues throughout the simulation (surface or core of the molecule). Although the disulfide bonds in the folded conotoxin can change their positions from the surface to the core of the molecule, they hold basic residues in a certain conformation on the surface. MD simulation of the peptides with one disulfide bond opened reveal stable structures and led to the assumption that these disulfide-deficient conotoxins maintain biological activity. This simplification would facilitate synthesis of these conotoxins to peptides with only four cysteine residues and also reduce the formation of possible isomers. A further reduction of the cysteine content would increase the flexibility of the structures resulting in less active compounds. Here, the combination of different theoretical approaches resulted in a promising tool for estimation of structural stability with respect to disulfide bridge removal prior to analog design and synthesis.

4.6 Conformational peptide isomers revisited: The impact of disulfide connectivity on structure and bioactivity

Full Paper

Authors

Pascal Heimer, Alesia A. Tietze, Enrico Leipold, Charlotte A. Bäuml, Jan-Pierre Schneider, Volker Schmidts, Christina M. Thiele, Desiree Kaufmann, Astrid Maass, Jan Hamaekers, Martin Hofmann-Apitius, Anja Resemann, Franz-Josef Mayer, Detlev Suckau, Oliver Ohlenschläger, Stefan H. Heinemann, Daniel Tietze and Diana Imhof

Submitted to Chemical Science, August 2017.

Preface

Although synthesis and characterization of 15 disulfide isomers of a μ -conotoxin is an extensive task, a detailed structural analysis of isomers with three disulfide bonds is of high interest. Selective disulfide bond formation using a specific protecting group strategy is challenging because of possible disulfide rearrangement during product formation. In general, the obtained product has only been characterized with one or two methods to elucidate the structure and disulfide connectivity, however, not every method can be used for analysis because accurate and unequivocal results cannot be provided. Herein, all 15 possible isomers of μ -conotoxin PIIIA have been synthesized and extensively analyzed. μ -PIIIA was used as model peptide because three different isomers could be isolated and characterized previously forming the basis for the present study.⁹⁶ For structure elucidation several methods namely HPLC, NMR and MS/MS were applied and the respective results were analyzed to compare the suitability for the distinction of all isomers. Furthermore, all peptides have been tested for their ability to block the skeletal ion channel Na_v1.4. To explain the mechanism of action a homology model was used to identify and characterize the toxin binding to the channel pore.

Conformational μ -PIIIA peptide isomers revisited: The impact of disulfide connectivity on structure and bioactivity

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

P. Heimer,^a A. A. Tietze,^b E. Leipold,^c D. Kaufmann,^d C. A. Bäuml,^a A. A. Paul George,^a A. Maass,^e J. Hamaekers,^e A. Resemann,^f F. J. Mayer,^f D. Suckau,^f O. Ohlenschläger,^g S. H. Heinemann,^c D. Tietze^{d*} and D. Imhof^{a*}

Peptides and proteins carrying a high number of cysteines can adopt various 3D structures depending on their disulfide connectivity. The unambiguous verification of such conformational isomers with more than two disulfide bonds is extremely challenging and experimental strategies for their unequivocal structural analysis are largely lacking. We synthesized the fifteen possible isomers of the 22mer conopeptide μ -PIIIA, applied 2D NMR spectroscopy and MS/MS for structure elucidation, and finally tested their activity to block the skeletal muscle voltage-gated sodium channel $\text{Na}_v1.4$. The activity profile of μ -PIIIA isomers was correlated with results of molecular docking investigations. Our study provides intriguing insights in how the disulfide connectivity alters the global fold of the toxin and thus influences bioactivity. We also show that analysis procedures involving a comprehensive combination of conventional methods are required for the unambiguous assignment of disulfides in cysteine-rich peptides and proteins and that chemical standard compounds are a crucial need for the structural analysis of such molecules.

Introduction

There are fundamentally different reports regarding the concomitant occurrence of conformational isomers of proteins and peptides with multiple disulfide bonds irrespective of their origin, i.e. isolation from biological material, synthetic production, or recombinant expression.^{1–6} Peptide and protein toxins were suggested to be suitable candidates for studying this particular issue because many of them are known to possess highly structured three-dimensional arrangements.^{1–9} In addition, different disulfide bond patterns of the same primary amino acid sequence were shown to result in different biological activities.^{10–13} In this respect, it should be noted that the importance of identifying the individual

disulfide connectivity is still underestimated, in particular in proteins harbouring more than four cysteine residues. At the same time it is commonly accepted that the elucidation of disulfide connectivity currently belongs to the major goals in proteomics and bioanalysis studies of peptide and protein therapeutics, such as monoclonal antibodies.¹⁴

For conotoxins derived from the venoms of marine cone snails already in 1991 Olivera *et al.* suggested degenerate ways to obtain congruent conformations.¹⁵ In view of recent findings,^{11,13} this can be interpreted as different disulfide connectivity might result in congruent conformations despite conserved residues and cysteine arrangement. This raises the following questions: i) How probable is the occurrence of individual or all possible disulfide isomers? ii) Which conformational isomers of disulfide-rich peptides and proteins are formed in commonly used production strategies? iii) Are the existing analytical methods and techniques sufficient to unequivocally identify individual disulfide isomers?

Irrespective of the research or development goals to be pursued, a detailed analytical characterization must be performed prior to research, clinical or commercial use. One of the main problems is chemical modifications of amino acids as a result of cellular posttranslational modification or of sample preparation and analysis processes causing product-related impurities such as undesired conformational disulfide isomers. We thus focused our attention on the 22mer conotoxin μ -PIIIA from *Conus purpurascens* containing six cysteine residues giving rise to three disulfide bonds, for which three differentially folded disulfide isomers were identified earlier in an oxidative self-folding approach.¹¹

^a Pharmaceutical Biochemistry and Bioanalytics, Pharmaceutical Institute, University of Bonn, An der Immenburg 4, D-53121 Bonn, Germany, E-mail: dimhof@uni-bonn.de

^b Clemens Schöpf Institute of Organic Chemistry and Biochemistry, Darmstadt University of Technology, Alarich-Weiss-Str. 4, D-64287 Darmstadt, Germany

^c Department of Biophysics, Center for Molecular Biomedicine, Friedrich Schiller University Jena and Jena University Hospital, Hans-Knöll-Str. 2, D-07745 Jena, Germany

^d Eduard Zintl Institute of Inorganic and Physical Chemistry, Darmstadt University of Technology, Alarich-Weiss-Str. 4, D-64287 Darmstadt, Germany, E-mail: tietze@chemie.tu-darmstadt.de

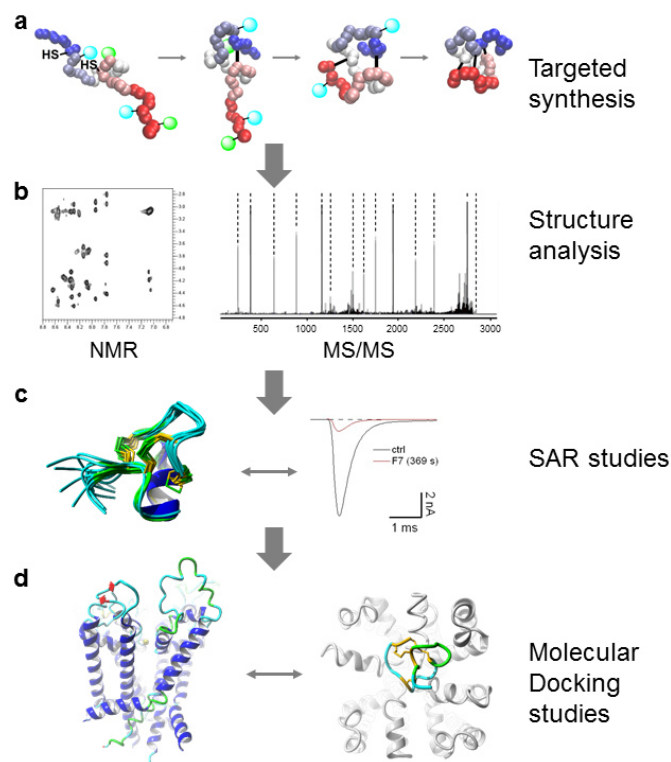
^e Fraunhofer Institute for Algorithms and Scientific Computing Department of Virtual Material Design, Schloss Birlinghoven, D-53754 Sankt Augustin, Germany

^f MALDI Applications and Proteomics R&D, Bruker Daltonics GmbH, Fahrenheitstr. 4, D-28359 Bremen, Germany

^g Leibniz Institute on Aging – Fritz Lipmann Institute, Beutenbergstr. 11, D-07745 Jena, Germany

† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

Here we describe the solid phase synthesis of the fifteen possible isomers using a targeted protecting group strategy for structural analysis and bioactivity testing (Scheme 1). We demonstrate that commonly used liquid chromatography methods are insufficient to identify distinct disulfide isomers while the disulfide connectivity of most - yet not all - isomers can be unambiguously assigned by applying a combination of MS/MS analysis and 2D NMR spectroscopy. All μ -PIIIA isomers were further analysed for their potency to block the skeletal muscle sodium channel $\text{Na}_v1.4$. In addition, interactions between isomers and $\text{Na}_v1.4$ channels were analysed on the molecular level by means of molecular docking simulations.



Scheme 1 Outline of the work process: synthesis, data acquisition, and analysis. (a) Targeted folding of linear μ -PIIIA to fully oxidized disulfide isomers in three steps according to a cysteine-specific protecting group strategy. (b) Structure elucidation was accomplished by combining solution NMR and high-resolution tandem mass spectrometry. (c) Electrophysiological studies revealed the intensity of blocking activity of the individual μ -PIIIA isomers at the voltage-gated sodium channel $\text{Na}_v1.4$. (d) Experimental results were correlated with results of molecular docking studies.

Results and discussion

Chemical analysis of synthetic μ -PIIIA disulfide isomers

Synthesis of the μ -PIIIA disulfide isomers was performed according to Fmoc solid-phase peptide synthesis with regioselective disulfide bond formation using the protecting group strategy trityl (Trt),

acetamidomethyl (Acm) and *tert*-butyl (tBu) for the respective pairs of cysteine residues (Supplementary Information, Table S1[†]).^{16,17} Peptides were obtained in high purity and acceptable yield using RP HPLC (Supplementary Information, Fig. S1, Table S2[†]). Purified oxidized peptides (Fig. S1, Table S2[†]) were submitted to structural analysis (MS/MS, NMR) and electrophysiological studies. Additional liquid chromatography experiments were performed in order to address the issue of clear distinction between highly similar conformational isomers with respect to their physico-chemical properties (Fig. S2[†]). After complete assignment of the disulfide connectivity (see below) equal amounts of three PIIIA isomers sharing one disulfide bridge, e.g. isomers **3**, **6**, **9** (Cys21-Cys22), isomers **3**, **4**, **5** (Cys4-Cys5), and isomers **13**, **14**, **15** (Cys4-Cys22), respectively, were mixed and re-injected for RP HPLC analysis using two different columns (C4 and C18) and two different instruments. These elution profiles revealed either a single peak or multiple peaks, but in none of the cases base-line separation of all three isomers was obtained (Fig. S2[†]).

Disulfide bond assignment of μ -PIIIA isomers by MS/MS analysis

MS/MS sequencing based on the analysis of suitable sets of precursor and fragment ion masses for an individual peptide was then performed to establish the disulfide bond pattern in all isomers (Supplementary Information). The most commonly applied peptide fragmentation techniques rely on qTOF and TOF/TOF mass analysers employing collision-induced decay (CID), in-source-decay (ISD), and post-source decay (PSD).^{18–20} Our intention to apply LC-ESI-qTOF-MS/MS and MALDI-TOF/TOF-MS/MS LID and MALDI MS/MS ISD in a direct comparative way, however, could not be realized due to incomplete fragmentation found in ESI MS/MS (Supplementary Information, Figures S3–S5, Tables S3, S4[†]). Enzymatic digestion was also not considered for μ -PIIIA isomers herein, because the cleavage sites result in unspecifically cleaved fragments.¹¹ In addition, the μ -PIIIA sequence includes adjacent cysteine residues in two regions (C4/C5 and C21/C22) (Fig. 1a), which is a characteristic structural feature of several peptides and proteins,^{21–28} however, additionally complicates analysis due to contradicting results for disulfide assignment.²⁹ Also, according to the applied synthetic strategy, a rearrangement of disulfide bonds might appear after cleavage of the tBu-protecting group and the following oxidation step in case of incomplete oxidation of the former Acm-protected Cys-pair.^{29–31} This earlier reported hypothesis^{29,30} is supported by the occurrence of several peaks in the crude peptide mixtures (Fig. S1[†]) and co-elution of these isomers (Fig. S2[†]) as aforementioned. It can be stated that for isomers **1**, **7**, **9**, and **11** no contamination with scrambled isomers was observed (see MS/MS and NMR results below), which can be explained also by the different retention times of the intended version and the possible impurities (Tables S2 and S3[†]). However, at least one other remaining isomer has the same retention time of the anticipated peptide and, thus, cannot be excluded by HPLC only (Supplementary Information).

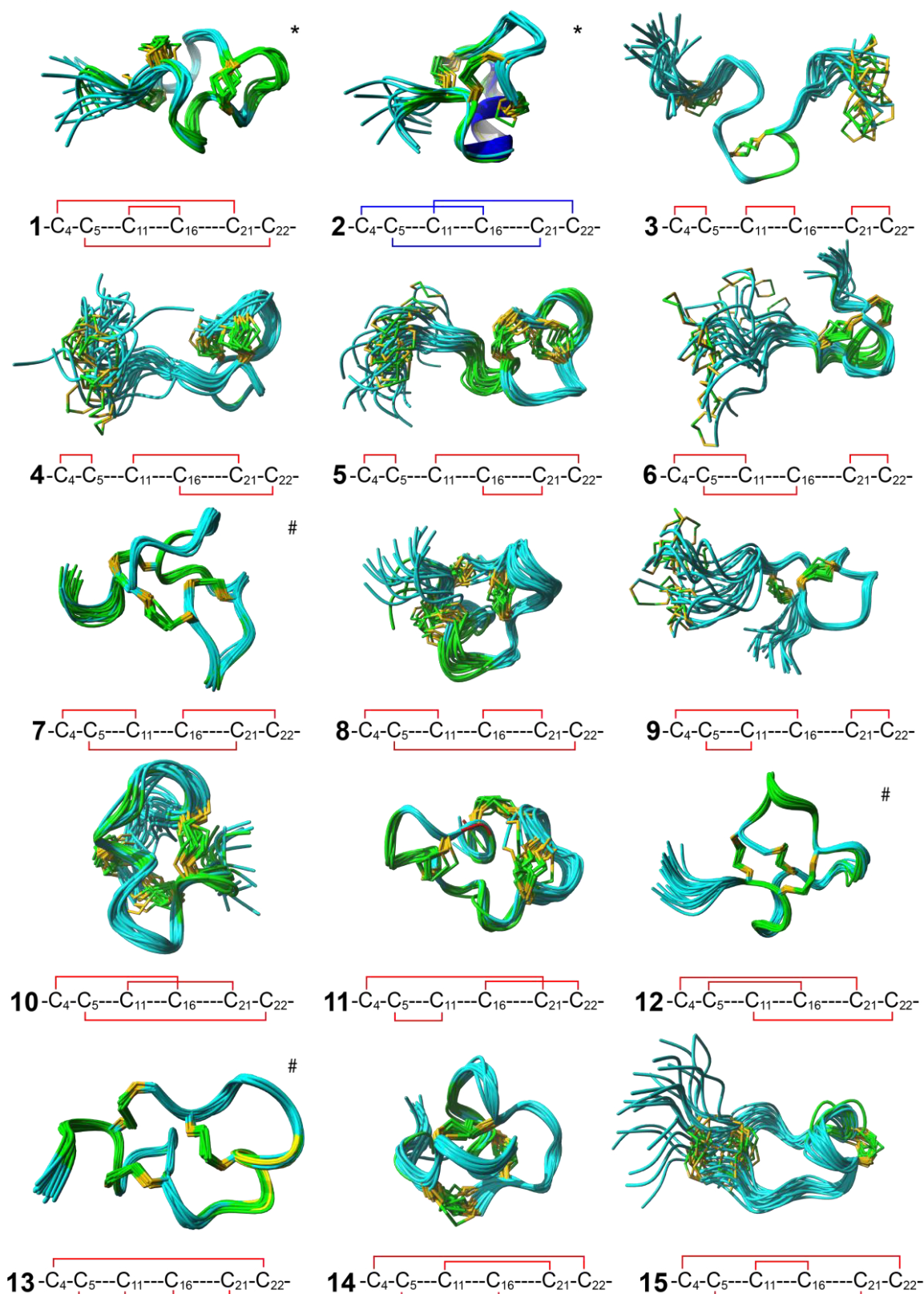
aZRLCC GFOKS CRSRQ CKOHR CC-NH₂**b**

Figure 1 (a) Sequence of μ -PiIIA and (b) NMR structures of all 15 μ -PiIIA isomers. For each isomer with the indicated disulfide-bond pattern, the solution NMR structures are presented as ensembles of 20 structures with the lowest energy. *Structures 1 and 2 were reported by us previously;¹¹ structure 2 represents the so-called native μ -PiIIA isomer.³³ #Structures predicted from MD simulation (see also Figure S6†).

To clarify the disulfide connectivity for each single isomer MS/MS analysis, subsequent NMR structure determination, and structure calculation for alternative isomers (Supplementary Information, Table S5†) were combined. MS/MS was performed with partially reduced and alkylated derivatives produced by an adapted earlier described protocol (Supplementary Information, Fig. 2, Figures S3–S5, Tables S3, S4†).³² In the MALDI-MS/MS LID TOF/TOF spectra internal fragments (IF) for respective 2- and 4-carbamidomethylated (CAM) species were taken into account, since distinct CAM IFs can only occur from a particular disulfide bridging in the oxidized peptide (Tables S3, S4†). Thus, it was possible to identify the same disulfide bonds in both the 2-CAM and 4-CAM species, which is indicative for successive bridge opening.

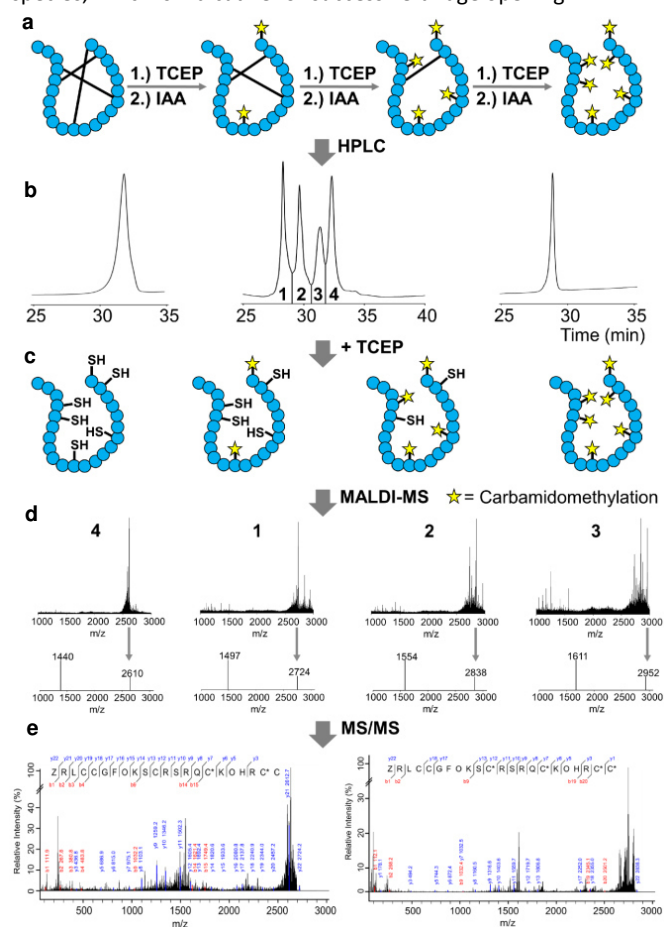


Fig. 2 Partial reduction and mass spectrometry analysis workflow. a) Schematic representation of partial reduction of the oxidized peptide isomer using TCEP and subsequent carbamidomethylation (yellow star) performed with iodoacetamide for different time periods. b) HPLC-elution profile of isomer 8 after three minutes of partial reduction and alkylation. All samples were fractionated by liquid chromatography prior to further characterization. c) After freeze-drying of collected fractions, reduction was completed to remove the remaining disulfide connections prior to MS analysis. d) MALDI-MS spectra of fractions obtained from isomer 11 and schematic representation of the respective internal fragment (IF) obtained from MS/MS fragmentation for distinct masses. This indicates the level and

position of carbamidomethylation within the peptide. The derivatization of the IF containing three cysteine residues is indicated by a molar mass of 1440 (no CAM), 1497 (one CAM), 1554 (two CAM), or 1611 (three CAM). Thereby different combinations between 2- and 4-times carbamidomethylated peptide and IF are possible, depending on the present disulfide pattern. e) Selected MALDI-MS/MS-spectra obtained from fraction 2 of isomer 5 with 2- and 4-times carbamidomethylated species after sequencing and manual evaluation including the obtained disulfide pattern (Table S3, Fig. S5†).

As a consequence and based on the combination of the results from MS/MS analysis and the synthetic strategy, we subsequently determined the 3D structure of the anticipated isomer, while also considering two possible alternative disulfide patterns during structure determination (Supplementary Information, Table S5†).

Structure analysis of μ -PIIIA isomers by 2D NMR

Solution structures of 12 μ -PIIIA isomers were determined by standard ^1H -based NMR spectroscopy (Fig. 1b). The sufficient spectral resolution of the 2D NMR spectra (TOCSY and NOESY/ROESY) at magnetic fields of either 14.1, 16.4 and/or 21.1 Tesla allowed for an unambiguous resonance assignment of almost all isomers, except for 7, 12, and 13, where the resulting NOE spectra were not evaluable. In addition, the NMR spectra revealed at least one minor toxin conformation to be present in select isomer solutions (with the exception of 1, 2, 11, 14, and 15) as indicated by an additional set of cross peaks. Unfortunately, the NMR data did not allow for structure determination of these less abundant isomer conformations. The structures of 7, 12, and 13 were modelled and the structural ensembles for these isomers were generated through a molecular dynamics simulation run (Fig. 1b, S6†).

For all structures a *trans* configuration of both hydroxyproline residues (Hyp8, Hyp18) was observed as dominant species and was identified by the occurrence of strong $\text{H}\delta$ - $\text{H}\alpha$ -1 and medium $\text{H}\delta$ - $\text{H}\beta$ -1 cross signals in case of Hyp8. Due to almost identical chemical shifts of $\text{H}\gamma$ and $\text{H}\alpha$ of Hyp18 and $\text{H}\alpha$ of Lys17 for most of the isomers only the medium intensity $\text{H}\delta$ - $\text{H}\beta$ -1 cross peak was observed, which indicates a *trans* configuration of Hyp18. No minor conformation comprising any of both Hyp residues in a *cis* configuration was observed, which is in contrast to the previous observations for μ -PIIIA.^{11,33}

The solution structure for each μ -PIIIA isomer was calculated using the respective NOE data (Fig. 1b, Supplementary Information, Tables S5–7†). The two alternative forms of disulfide connectivity (Table S5†) from the synthetic strategy and Poppe's method³⁴ for disulfide bond assignment were considered to test the unambiguity of the experimental NMR restraints for the structure calculations. In contrast to our previous studies of three μ -PIIIA isomers obtained from an oxidative self-folding approach,¹¹ the solution structure of isomer 3 could be obtained herein due to higher purity of the sample yielding high-resolution NOE data for structure calculation.

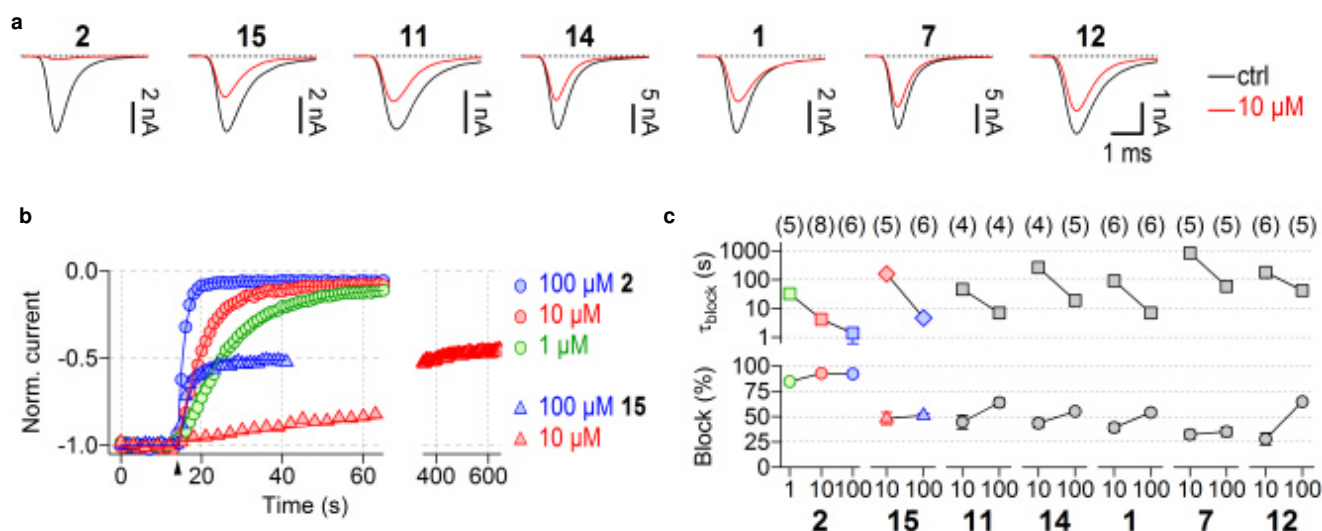


Fig. 3 Disulfide isomers of μ -PIIIA differentially inhibit $\text{Na}_v1.4$ -mediated currents. a) Representative current traces of transiently expressed $\text{Na}_v1.4$ channels evoked at a test potential of -20 mV before (black, ctrl) and after (red) application of $10 \mu\text{M}$ of the indicated μ -PIIIA isomers. b) Normalized peak current amplitudes obtained from repetitively evoked current responses were plotted as a function of time to follow the time course of current block mediated by various concentrations of μ -PIIIA isomers **2** and **15**. Continuous lines are single exponential data fits used to characterize the onset of channel block. The arrowhead marks the start of μ -PIIIA application. The time axis was split to illustrate that channel block by isomer **15** saturates at about 50% suggesting that isomer **15** seals the channel pore only partially. In contrast, isomer **2**-mediated channel inhibition saturates at about 95% with an IC_{50} of 105.3 ± 29.9 nM. c) Steady-state block (bottom) as well as the associated single exponential time constant (top), describing the onset of channel block, for various concentrations of the indicated isomers. Lines connect data points for clarity. Symbols and color-coding of data obtained with isomers **2** and **15** are as in b). Numbers of individual experiments, n , are provided in parentheses.

The resulting structures fall into three categories with respect to their resolution, which is reflected by the overall deviation of the backbone coordinates of the structural ensemble (RMSD, root mean square deviation). The structures are highly dissimilar, ranging from a compact fold (isomers **2**, **11**, **14**), an intermediate fold of compact and flexible parts (isomers **1**, **8**, **10**) to very flexible, open structures as found for isomers **3-6**, **9** and **15** (Fig. 1b). Nevertheless, all structures revealed similar backbone coordinates for the loop between Cys11 and Cys16 with an average backbone deviation of about 1.4 \AA .

The NMR structural ensembles of isomers **2**,¹¹ **11**, and **14** exhibited the highest resolution (RMSD 1.22 , 0.94 , 0.90 \AA), followed by isomers **1**, **8**, and **10** with a medium resolution (RMSD 1.46 , 1.48 , 1.54 \AA). Isomers **3-6**, **9**, and **15** generated only low-resolution structures (RMSD ranging from 1.99 to 2.77 \AA , Fig. 1, Table S5[†]). The low resolution of the NMR ensembles obtained for isomers **3-6**, and **9** are a direct consequence of their flexible N- and/or C-terminal portions, i.e. if C4-C5 (isomers **3-5**) and/or C21-C22 are connected (isomers **3**, **6**, **9**). However, the remaining part of the molecule shows a low backbone deviation and thus is rigid (Fig. 1b, Table S5[†]). In contrast, the compact fold and the low global backbone deviation of isomers **2**, **11**, and **14** are indicative of a very rigid structure. Despite their rather compact fold, isomers **1**, **8**, and **10** exhibit a much higher structural inhomogeneity than the rigid ones, probably due to their higher global flexibility. Apart from that, isomer **15** seems to stand out regarding its highly inhomogeneous backbone conformation (Fig. 1b), which might be caused by its unusual, ladder-like disulfide bond pattern connecting C4-C22, C5-C21, and C11-C16.

Calculations of the aforementioned alternative structures were performed considering two different disulfide linkages (Table S5[†]), which were also carried out for the earlier described isomers **1** and **2**.¹¹ Reasonable alternative structures were obtained for isomers **1**, **2**, **10**, **14**, and to a lesser extent also for isomers **4**, **5**, and **8** (Table S5[†]) applying the individual sets of experimental NMR restraints used to calculate the structures. However, according to MS/MS analysis we could confirm the intended disulfide connectivity in all of these isomers.

Bioactivity of μ -PIIIA isomers at $\text{Na}_v1.4$ ion channels

The most likely biological function of μ -PIIIA is to act as a neurotoxin that blocks the permeation pathway of voltage-gated Na^+ (Na_v) channels. Previous reports have shown that μ -PIIIA has the highest potency to interfere with skeletal muscle channels ($\text{Na}_v1.4$) although also being active on neuronal Na_v channels and even select K_v channels.³⁵ We therefore applied all isomers to HEK 293 cells expressing human $\text{Na}_v1.4$ channels and measured the impact of the toxin isomers on depolarization-elicited Na^+ inward currents, expressed as the time course of current block and the maximal block after long toxin exposure (Fig. 3, Fig. S7[†]). Figure 3 summarizes the inhibitory activity of the most active isomers **1**, **2**, **7**, **11**, **12**, **14**, and **15**. As demonstrated in Figure 3a, currents mediated by human $\text{Na}_v1.4$ were inhibited by all isomers, albeit to different degrees. At a concentration of $10 \mu\text{M}$, isomer **2**, which has the native disulfide connectivity of μ -PIIIA, was most effective in inhibiting $\text{Na}_v1.4$ channels, followed by isomers **15**, **11**, **14**, **1**, **7**, and **12**. Analysis of the concentration dependence revealed that isomer

2 blocked human Na_v1.4 with an apparent IC₅₀ of 105.3 ± 29.9 nM, which is comparable to the value for the paralog rat Na_v1.4 channels (103.2 ± 9.9 nM).¹¹

As observed for some μ -conotoxins, even at a saturating concentration of 100 μ M a small current component of 7.2 ± 2.2% remained, indicating that channel occupancy by isomer **2** does not completely eliminate Na⁺ conduction (Fig. 3a, b). The remaining isomers (**15**, **11**, **14**, **1**, **7**, **12**) blocked Na_v1.4 channels less potently than isomer **2**. In addition, onset of block as estimated with single-exponential functions was substantially slower than for isomer **2**, thus precluding faithful assessment of channel block at lower concentrations than 10 μ M and the associated IC₅₀ values (Fig. 3c). Assessment of higher concentrations revealed that particularly isomers **15** and **7** are interesting: total current block at 10 and 100 μ M were virtually identical (**15**: 48.2 ± 5.9% and 51.1 ± 2.2%, respectively; **7**: 32.3 ± 4.3 and 35.1 ± 4.6, respectively), while the time constant characterizing the kinetics of onset of block, τ_{block} , roughly scaled linearly with the concentration (**15**: 10 μ M: 160.0 ± 34.9 ms; 100 μ M: 47.0 ± 9.7 ms; **7**: 10 μ M: 853.0 ± 195.0 ms; 100 μ M: 59.0 ± 18.1 ms; Fig. 3c). This result indicates saturated association between isomers **15** and **7** and the channel, with imperfect (about 50% and 35%, respectively) occlusion of the Na⁺ permeation pathway. For isomers **11**, **14**, **1**, and **12** saturation of channel block at 100 μ M peptide concentration was not apparent (Fig. 3c). The channel block of isomers with even lower potency (**3**–**6**, **8**, **10**, and **13**) were analysed at 10 μ M. As shown in Figure S7†, sample **13** was most active among this group, as it inhibited Na_v1.4 by 24.6 ± 3.0% followed by samples **4** (21.7 ± 3.5), **9** (17.7 ± 1.7%), and **3** (17.6 ± 3.6%). Isomer samples **6**, **8**, **5**, and **10** reduced Na_v1.4-mediated currents by less than 15% (**6**: 12.6 ± 1.3%, **8**: 11.4 ± 2.2%, **5**: 6.7 ± 2.1%, **10**: 1.0 ± 3.0%), and thus were considered inactive under these conditions.

In silico toxin binding studies

Interactions of the μ -PIIIA isomers with the Na_v1.4 channel were further investigated employing a homology model which comprised the central pore module of the channel similar to the model developed earlier by us¹¹ and others.³⁶ We identified the recently published structure of the voltage-gated calcium channel Ca_v1.1 (PDB ID 3jbr³⁷) as a suitable template, allowing to model the large, pore-surrounding, extracellular loops that contain important interaction sites for μ -conotoxins, which were not yet included in available homology models (Supplementary Information, Fig. S8†). The recently solved structure of a eukaryotic voltage-gated Na⁺ channel (Na_vPaS, pdb ID 5X0M³⁸) was not available, when this study was started. However, our homology model and the structure of the central pore module of Na_vPaS are in good agreement, deviating by only 1.9 Å (Fig. S9b†).

All 15 μ -PIIIA isomers (Fig. 1) were docked to the Na_v1.4 homology model using the HADDOCK easy web interface.^{39,40} The docking results were clustered according to the interface-ligand RMSDs and scored corresponding to the HADDOCK scoring function (Table S8†). The best scoring channel–toxin complex of the highest

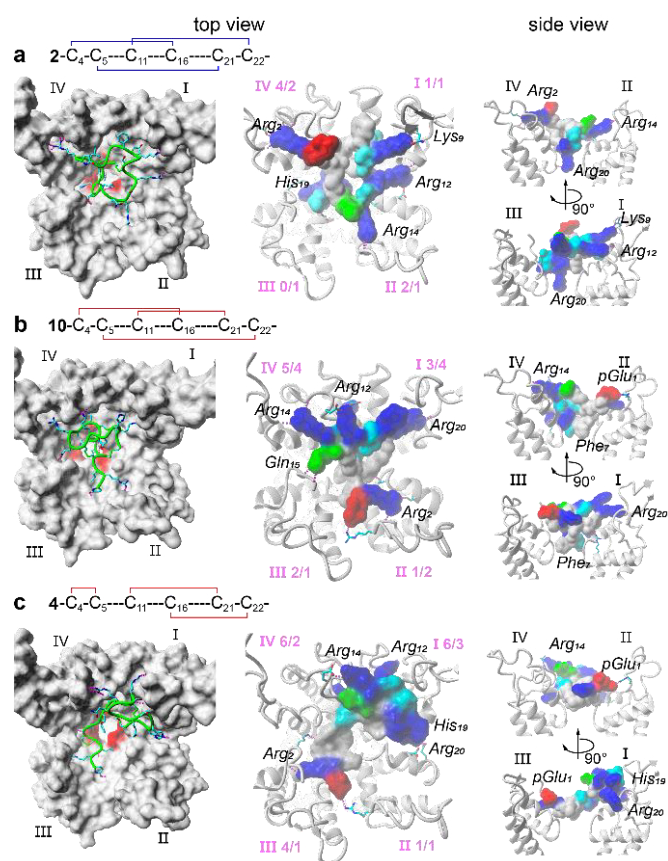


Fig. 4 a – c) Visualization of the best scoring μ -PIIIA–Na_v1.4 complex conformations obtained from docking experiments. The four Na_v1.4 domains are indicated. Left panel – top view of the toxin-channel complex. The Na_v1.4 channel surface (Van der Waals) is given in gray, the selectivity filter motif (DEKA) is highlighted in red. The toxin is shown in green with side chain atoms present (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, sulfur – green). Middle and right panel – top and side view of the toxin-channel complex. The toxin is shown in surface representation (Van der Waals) and colored according to amino acid properties (coloring scheme: basic – blue, acidic – red, polar – cyan, non-polar – gray). The number of hydrogen bonds formed by the toxin with each channel domain (loops/SS1-SS2) is highlighted in purple distinguishing between hydrogen bonds formed with the TM bridging loops (S5-SS1, SS2-S6) and the pore helices (SS1-SS2).

scoring cluster was subjected to further analysis. In all complexes evaluated, μ -PIIIA occupied the central pore interface. Viewing the complexes from the top revealed that the binding mode of the isomers is further distinguished mainly by the orientation of their arginine residues (Arg2, Arg12, Arg14, Arg20) and the positioning of their lysine residues (Lys9, Lys17) or Phe7 and Pyr1 with respect to the channel's surface (Fig. 4, Fig. S10†). We observed an inhomogeneous one-way orientation of the majority of the arginine residues in the direction of the pore loops connecting the transmembrane segments S5 with SS1 and SS2 with S6 of domains I and IV (Fig. S8†) in isomers **4**, **10** (Fig. 4b, c), **5**, **6**, and **15** (Fig. S10e, g, c†). Among these four isomers, isomers **4**, **5**, and **10** showed a separate positioning of Arg2 together with Pyr1 offside the other arginine residues in the direction of the pore domains III and II (Fig. 4b, c and Fig. S10e†). In contrast, all the remaining isomers (**1**–**3**, **7**–

9, 11-14) exhibit a rather uniform distribution of their basic Arg residues across the channel's pore entry (Fig. 4, Fig. S10[†]).

The docking results further suggest that the acidic residues in the S5-SS1 and SS2-S6 bridging pore loops serve as anchor points for the basic toxin residues allowing the toxin to span over the central pore. Moreover, the high number of acidic residues within these loop regions (Fig. S8[†]) might indicate a key role for toxin affinity. In addition to a versatile arginine orientation, an arginine residue protrudes more deeply into the pore and thus establishes additional interactions with the selectivity filter region such as observed for Arg20 of isomers **1** and **2** (Fig. 4a, Fig. S10a, viewing the complexes in side view[†]). However, such arginine orientation might be beneficial for an efficient pore block, but may not be critical as revealed by the bound conformation of **11**, **14** and **15** (Fig. 4b, Fig. S10l,d, c[†]).

Interestingly, the docking result of the native isomer **2** is distinct from previous studies. Our model, however, confirms mutation analysis for native μ -PIIIA,^{35,41} since all basic residues are involved in hydrogen bonding being almost homogeneously distributed among the channel surface and thus contributing to the effective pore block. This also supports the fact that PIIIA lacks a critical single residue for an all-or-none inhibition of the current flow. This could not be fully explained by previous models, in which the pore loops were not considered in the respective homology models.^{36,41-43} The bound conformation of **2** is also in agreement with the importance of distinct PIIIA residues, e.g. Arg14 and Arg2,^{35,41} even though Arg14 does not protrude into the channel pore (Fig. 4a). In our model, μ -PIIIA covered the pore slightly asymmetric, showing more hydrogen bonds to domain I and IV (8 in total, Fig. 4a) than to II and III (4 in total, Fig. 4a, Table S9[†]). Besides, Arg14, which is hydrogen-bonded to the S5-SS1 loop of Na_v1.4 domain II, might provide an important anchor to effectively occlude the pore being lost upon mutation. Our data also indicate that the pore loops are involved in toxin binding and should therefore be considered in future *in silico* binding studies.

Experimental Section

Materials and Methods

Peptide synthesis. The amino acid sequence of μ -PIIIA was automatically assembled using Fluorenylmethyloxycarbonyl (Fmoc) chemistry and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), N-methylmorpholine/DMF (1:1) couplings on a Rink-amide resin (0.28 mmol/g). Protecting groups of amino acids other than cysteine were as follows: Boc-Pyr, Arg(Pbf), Hyp(tBu), Lys(Boc), Ser(tBu), Gln(Trt), and His(Trt). Synthesis of the linear peptide precursor was performed using Trt-, Ac- and tBu-protected cysteine residues according to the intended disulfide connectivity (Table S1). The choice of the synthesis strategy was guided by the idea of combining isomers in which amino acid sequence including protecting groups is identical up to a distinct residue, and then the resin was split into respective parts to enable assembly of the remaining residues. Partially Ac- and tBu-protected linear peptides were obtained in high purity after

cleavage with TFA and a scavenger mixture (82.5% trifluoroacetic acid (TFA), 5% water, 5% phenol, 5% thioanisole, 2.5% ethanedithiol) for 3 h at room temperature. After MS characterization the purified linear semiprotected peptides were subjected to a stepwise oxidation protocol A - C (see below). For the formation of the first disulfide bridge (protocol A), the linear purified peptide was dissolved in a mixture of isopropanol/water (1:2) with a final peptide concentration of 0.05 mM and oxidation on air under basic conditions (pH 8.5 adjusted with 1 M NaOH). The reaction was monitored via HPLC and MS following iodoacetamide (IAA) derivatization until complete oxidation was confirmed. After lyophilization of the precursor, the crude powder was used without further purification for the next deprotection step of Ac- groups followed by the oxidation step (protocol B). Therefore, a solution containing peptide (final concentration 0.05 mM) in a mixture of isopropanol/water/1 M HCl (80:12.5:7.5) was treated with iodine (final concentration 0.15 mM) in methanol at room temperature for 3-52 h, i.e. until completed. The reaction was stopped using a 1-M solution of ascorbic acid in a 5-fold excess relative to iodine. Monitoring and confirmation of oxidation was performed with HPLC and MS, respectively. Following lyophilization of the crude reaction mixture, the final deprotection of the tBu-protecting groups was performed (protocol C). The crude material was dissolved in TFA corresponding to the amount of peptide (i.e. for 1 μ mol of peptide 1 mL TFA was used), and containing a scavenger cocktail consisting of diphenylsulfoxide/anisole/trichloromethylsilane (1:10:15). The mixture was gently shaken for 3-5 h at room temperature and monitored via HPLC and MS. The individual peptide was then precipitated in cold diethyl ether, centrifuged and washed at least four times with diethyl ether prior to freeze-drying from 80% tert-butanol/water.

Peptide purification. The crude peptides were purified by semipreparative reversed-phase HPLC using a Shimadzu LC-8A system equipped with a C18 column (Knauer Eurospher 100). The gradient elution system was 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile/water (9:1) (eluent B). The peptides were eluted with a gradient of 0-50% eluent B in 120 min and a flow rate of 10 ml/min. The peaks were detected at 220 nm. Collected fractions were combined, freeze-dried and stored at -20 °C.

Peptide chemical analysis. Peptide purity after separation was confirmed with a Shimadzu LC-10AT chromatograph equipped with a Vydac 218TP54 column (C18, 5 mm particle size, 300 Å pore size, 4.6 x 25 mm) and a Vydac 208TP54 column (C8, 5 mm particle size, 300 Å pore size, 4.6 x 25 mm). The peptides were eluted with a gradient of 10-40% eluent B in 30 min. The flow rate was 1 ml/min, eluent A was 0.1% TFA in water, and eluent B was 0.1% TFA in acetonitrile; detection was at 220 nm. The existence of linear semiprotected peptide carrying Ac- and tBu protecting groups as well as each oxidation step after their removal was confirmed by MALDI-TOF mass spectrometry on an autoflex Bruker mass spectrometer (Bruker Daltonics GmbH, Bremen) using α -cyano-4-hydroxycinnamic acid as matrix. The reflector mode was used for

measurements and Peptide Calibration Standard II for molar masses below 6.000 g/mol. Derivatization of the observed samples with iodoacetamide confirmed the absence of free thiol groups, i.e. the formation of the respective disulfide bridge (Figure S5). For co-elution of different isomers the respective mixtures were measured using the aforementioned Shimadzu LC-10AT chromatograph equipped with a Vydac 218TP54 column (C18, 5 mm particle size, 300 Å pore size, 4.6 x 25 mm) and a Waters 600 HPLC system equipped with a MultoHigh U C4 column (CS - Chromatographie Service GmbH, 2 µm particle size, 120 Å pore size, 4 x 100 mm). Peptide isomers were eluted with a gradient of 5-50% eluent B in 20 min. Flow rate and eluent system were identical to the method described above; detection was at 214 nm.

Amino acid composition of the peptides was determined by amino acid analysis using a LC 3000 system (Eppendorf-Biotronik GmbH, Hamburg). Hydrolysis was performed with 6 N HCl in sealed tubes at 110 °C for 24 h. The obtained analyses supported the expected quantitative results. Amino acid analysis in combination with HPLC analysis (calibration curve) was used for the determination of peptide concentrations in solution prior to the further experiments to ensure equal concentrations in analysis and biological testing.

MS/MS analysis of disulfide connectivity. The partial reduction and alkylation protocol was adapted from an earlier study.⁴⁴ The pure peptide was dissolved in citrate buffer (0.05 M, pH 3.0) containing 0.02 M tris(2-carboxyethyl)phosphine (TCEP). The final peptide concentration was approximately 100-300 µg/ml. The mixture was then incubated at room temperature, and samples were taken at different times ranging from 0 s to 30 min. The samples were mixed with three equivalents of alkylation buffer (0.5 M Tris-acetate pH 8.0, 2 mM EDTA and 1.1 M iodoacetamide) to stop the reaction and perform carbamidomethylation of the free thiol groups from partial reduction. The alkylation reaction was terminated by addition of one volume of 10% TFA after 5 min. Samples were stored on dry ice until further analysis with HPLC using the following gradient: 10% of eluent B for 15 min, then linear increase of eluent B up to 35% for another 25 min. Fractions were collected and further subjected to MS/MS sequencing. Linear and fully carbamidomethylated µ-PIIIA were subjected to sequencing on an ultrafleXtreme MALDI TOF/TOF mass spectrometer or an ultraflex III MALDI TOF/TOF mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) and were analyzed using flexAnalysis 3.4 and BioTools 3.2, respectively. For MS/MS analysis of the disulfide connectivity in the collected fractions from HPLC separations (Figure S2), the samples were measured in the oxidized and reduced form as follows. The peptides (10 - 100 µg in 10 - 50 µl water containing 0.1% TFA) were reduced by adding the appropriate volume of a 100 mM aqueous TCEP solution to a final concentration of 10 mM TCEP, and incubation was carried out for 1 h at 37 °C prior to further analysis by tandem mass spectrometry.

A representative ESI-MS/MS spectrum of linear µ-PIIIA is shown in Figure S4 from an LC-ESI micrOTOF-Q III mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) coupled with a Dionex Ultimate 3000 (Thermo Scientific, Dreieich, Germany) equipped

with a EC100/2 Nucleoshell RP18 Gravity 2.7 µm column (Macherey-Nagel, Düren, Germany). Analysis of the MS data was performed using Bruker Compass Data Analysis 4.1. LC conditions used were as follows: Eluent A was water with 0.1% acetic acid, while eluent B was acetonitrile containing 0.1% acetic acid. A gradient of 0% -60% of eluent B in 12 min was used and detection was at 220 nm.

Since fragmentation of the linear peptide was insufficient even after optimization of MS/MS parameters (Figure S4), no further analysis of the disulfide connectivity of the different carbamidomethylated species was performed with this method. MS/MS analysis and sequencing was carried out using an ultrafleXtreme MALDI TOF/TOF mass spectrometer or an ultraflex III MALDI TOF/TOF mass spectrometer (both Bruker Daltonics GmbH, Bremen, Germany) (Table S3). Different techniques were applied for preparation and analysis of the samples. Samples were spotted on ground steel or AnchorChipTM targets using saturated solution of HCCA (α-cyano-4-hydroxycinnamic acid) for ground steel and 0.7mg/ml for AnchorChipTM targets or sDHB (90:10 mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid) in 50% acetonitrile/water containing 0.1% TFA as matrices. Initially, both ISD and LID TOF/TOF spectra were acquired. Since several samples contained both 2- and 4-times carbamidomethylated species, MS/MS LID measurements were performed in order to preselect the parent ion for fragmentation. Although ISD experiments generally revealed more fragments, especially regarding y- and c- ion series of the whole sequence, a greater amount of substance is required, pmol rather than fmol on target. Furthermore, ISD was not suitable for analysis of mixtures containing 2- and 4-carbamidomethylated species in one sample due to overlapping fragments resulting from both species.

Peptide Calibration Standard II (Bruker Daltonics GmbH, Bremen, Germany) or ubiquitin (Sigma-Aldrich) were used for calibration of the respective MALDI spectra. In some cases further sample purification was necessary after complete reduction with TCEP. C18-concentration filters (ZipTip[®]) were used to remove TCEP from the reduced samples according to the annexed protocol from Millipore (Merck Millipore, Molsheim, Germany). Data processing and evaluation was performed using the software packages Bruker Compass Data Analysis 4.1, flexAnalysis 3.4, and BioTools 3.2.

NMR experiments and structure analysis. NMR spectra were recorded for µ-PIIIA isomers in 90% H₂O/10% D₂O using the freeze-dried solid compound at 293 K on Bruker Avance III spectrometers with proton frequencies of 600, 700 or 900 MHz equipped with cryogenic probes. In all NMR spectra, the ¹H peak from water was used as a chemical shift reference by setting its frequency at 4.7 ppm. All NMR data were processed and analyzed using TopSpin 2.1 (Bruker) and CcpNMR Analysis⁴⁵. The proton resonance assignment was performed by a combination of 2D [¹H,¹H]-DQF-COSY, [¹H,¹H]-TOCSY, [¹H,¹H]-NOESY and [¹H,¹³C]-HSQC spectra using water suppression. Distance constraints were extracted from [¹H,¹H]-NOESY spectrum acquired with a mixing time of 120 ms and a recycle delay time of 1.5 s. Upper limit distance constraints were

calibrated according to their intensity in the NOESY spectrum and the intensity of geminal protons was used for peak intensity calibration. Torsion angle constraints were obtained from ^1H and ^{13}C chemical shift analysis using DANGLE.⁴⁶ Structure calculations and refinements were performed using the respective disulfide connectivity as additional restraints with YASARA structure.^{47–50} The 20 structures with the lowest energy were selected to represent the NMR solution structures (Figure S7, Table S6 and S7).

Electrophysiological experiments. Human SCN4A (encoding the Nav1.4 channel α subunit, UniProt ID P35499) on a plasmid with CMV promoter was transiently expressed in HEK 293 cells as shown previously.¹¹ Co-transfection of a plasmid encoding the CD8 antigen ensured visual detection of transfected cells with CD8-specific Dynabeads (Deutsche Dynal, Hamburg, Germany). Currents were measured with the whole-cell patch clamp method 24–48 h after transfection.¹¹ The patch pipettes contained (in mM): 35 NaCl, 105 CsF, 10 EGTA (ethylene glycol bis(2-amino-ethylether)tetraacetic acid), 10 HEPES (pH 7.4 with CsOH). The bath solution contained (in mM): 150 NaCl, 2 KCl, 1.5 CaCl₂, 1 MgCl₂, 10 HEPES (pH 7.4 with NaOH). Holding potential was -120 mV, Na⁺ currents were elicited with depolarizing steps to $+20$ mV. Series resistance was corrected electronically up to 80%. Peptides, diluted in the bath solution, were applied focally to cells under consideration with a fine-tipped glass capillary. Time course of peak current decrease after peptide application was described with single-exponential functions.

Homology modelling, docking and simulation. The YASARA molecular modeling program was used to build the homology model of the pore module of the Nav1.4 channel and for molecular dynamics simulations.^{48,49} The crystal structure of Cav1.1 (PDB ID 3jbr³⁷) was used as template to model the Nav1.4 sodium channel which has an overall sequence identity of 21 %. Loop regions which could potentially be part of the toxin binding epitope were modeled or partly modeled with respect to their sequence conservation score (Figure S9)³⁷ and according to the UniProt sequence alignment⁵¹ of Nav1.4 (UniProt ID P35499) and the template sequence (UniProt ID P07293, Figure S9). In order to correctly identify transmembrane segments and loop-regions, the crystal structure of the bacterial sodium channel NavAB (PDB ID 3rvy⁵²) was structurally aligned with the template structure of Cav1.1. Structural alignments were achieved with the MUSTANG algorithm.⁵³ If necessary, loops were extended by insertion of the respective residues into the template loop sequence by Yasara's build loop command (Figure S9).⁵⁴ In case loops had to be shortened, the respective residues were deleted and the resulting gap was closed. Remodeled loops were energy minimized in vacuum employing the NOVA⁴⁹ force field while the rest of the protein was kept fixed using the LINCS method.⁵⁵ The crude homology model was then energy minimized, refined through an unrestrained all-atom molecular dynamics simulation for 2 ns in a membrane environment (PEA) and explicit water, and energy

minimized again prior to the docking studies using the Amber14 force field.^{56,57}

The μ -PIIIA isomers were docked individually to the constructed homology model of the Nav1.4 channel. Therefore, as a preparatory step, we selected all the specific residues known to be located on the channel's surface and thus depicting the active, toxin-interacting area as outlined in Figure 4. To perform the docking, the Easy Interface of the HADDOCK web service^{39,40} was used, which in turn requires the definition of active residues that are directly involved in the ligand interaction. The individual indices of the previously selected set of interacting residues were given as input. For the toxin itself, all 22 residues were designated to be involved in the interaction and thus set as active.

For subsequent analyses, the best-scored HADDOCK cluster from each isomer's docking run was extracted, i.e. the one assigned to the lowest Z-score.

To obtain a rough estimate of the μ -PIIIA binding affinity, the best scoring HADDOCK cluster was energy minimized employing the Yasara2⁴⁸ force field and re-scored using VINA's scoring function employing the YASARA structure software package for set-up purposes keeping the ligand flexible and the receptor fully rigid. The individual Z-scores are listed in Table S8 together with the corresponding binding energies resulting from the VINA re-scorings. Structure predictions of isomers **12** and **13** were performed using the lowest energy structure of the alternative NMR structure calculation of isomer **1** (for **12**) and **14** (for **13**) as a starting point. The starting structure of isomer **7** was derived from the conformation of isomer **8**, which had a similar disulfide bond pattern. The respective structures were then subjected to a 5 ns molecular dynamic simulation at 398 K followed by a 45 ns production run at 298 K, employing the Yasara2 force field. The structural ensembles were generated from 20 snapshot collected every 1 ns during the last 20 ns of the MD run.

If not otherwise stated, energy minimizations and molecular dynamics simulations were performed in explicit water (TIP3P water model⁵⁸) using the PME method⁵⁹ to describe long-range electrostatics at a cut-off distance of 8 Å (for Yasara2 and Amber14 force field) or 10.5 Å (for NOVA force field) at physiological conditions (0.9% NaCl, 298 K, pH 7.4⁵⁰). The energy minimization was achieved by a steepest descent minimization followed by a simulated annealing minimization until convergence (<0.05 kJ/mol/200/steps). Charged amino acids were assigned according to the predicted pKa of the amino acid side chains by Ewald summation⁵⁰ and were neutralized by adding counter ions (NaCl).

Molecular graphics were created with YASARA⁶⁰ (www.yasara.org) and POV-Ray (www.povray.org).

Conclusions

In conclusion, we have described synthesis and in-depth analysis of the 15 possible μ -PIIIA isomers with 12 representing new molecules and three being reproduced from earlier studies.¹¹ We show that a combination of all data obtained from HPLC, MS/MS fragmentation, and NMR analysis were required to unambiguously identify the

disulfide connectivity of each individual μ -PIIIA isomer. MALDI-MS/MS sequencing facilitated the identification of the individual isomers after partial reduction and derivatisation. In this way, it could be demonstrated that the isomers **1-5**, **7-11**, and **13-15** did not contain any other peptide version with an alternative fold. The respective disulfide connectivity was also confirmed for **6** and **12**, but these isomers revealed contaminations with scrambled variants in MS/MS analysis. The 3D structures of the μ -PIIIA isomers, represented by their respective NMR ensembles, strongly suggest that the global fold of the toxin is a direct consequence of the disulfide bond pattern. Surprisingly, the resulting structures are not only different in their global conformation but also in their backbone flexibilities. As one would expect, connecting the N- and/or C-terminal cysteines C4-C5 (**3-5**) and/or C21-C22 (**3**, **6**, **9**) resulted in flexible terminal regions. All other disulfide bond patterns resulted in a more compact fold, such as the native isomer **2**. Considering the block efficacy (Fig. 3, Fig. S7†) and the structure of each isomer (Fig. 1) it became evident that only a few disulfide combinations are compatible with an efficient inhibition of Na_v1.4 channels. The most active isomers **1**, **2**, **7**, **11**, **12**, **13**, **14**, and **15** all contain at least one disulfide bond that connects the N- and C-terminal sections of μ -PIIIA, thereby largely reducing the overall structural flexibility of these isomers. In contrast, most isomers with strongly reduced (**3**, **4**, **9**) or no activity (**5**, **6**) comprise highly flexible N- or C-terminal portions due to a connection between consecutive cysteines. This suggests that structurally rigid μ -PIIIA isomers are generally more potent to inhibit Na_v1.4 channels compared with isomers with more flexible structures. However, the almost inactive isomers **8** and **10** do not fit to this simplified interpretation because both contain a disulfide bond linking their N- and C-terminal parts. Interestingly, the NMR structural ensembles of both isomers showed a rather high structural inhomogeneity compared to the active isomers, which might serve as a possible explanation of their poor channel block. The docking results of both isomers revealed that Arg2 enforces the toxin to span over the pore (Fig. 4b, Fig. S10h†). Due to the local flexibility of the N-terminus, the resulting interaction of Arg2 with the flexible channel loop might be rather low, thus partly explaining the low bioactivity of **8** and **10**. Our data suggest that the channel selects for a toxin conformation indicated by the higher block rates of the more rigid toxin isomers. This clearly contradicts previous modeling data for μ -PIIIA,^{36,41–43} which suggested multiple toxin conformations to bind to and block the central pore. This observation might be explained by the missing loop regions of their Na_v1.4 models, which are indicated as being important to span the toxin over the central pore in our model.

Thus far, systematic studies of the structural and functional impact of the disulfide bonding in all possible variants have been confined to peptides with two disulfide bridges only.^{10,13} This study on μ -conotoxin PIIIA was inspired by the existence of numerous natural peptides and proteins with important biological function or high therapeutic potential that possess at least three disulfide bonds and one pair of adjacent cysteines. Typical examples are defensins, hepcidin, insulin-like growth factor I, norrin, placental

growth factor, resistin, sarcospan, vascular endothelial growth factor A, and Kunitz-type trypsin inhibitor SdPI.⁶¹ Our results for μ -PIIIA isomers revealed the complexity of synthesis, chemical analysis, and structural characterization of such conformational isomers and the urgent need for standards for elucidating the disulfide connectivity and 3D structure of disulfide-rich peptides and proteins.

Acknowledgements

We like to thank V. Schmidts, C. Thiele (TU Darmstadt), M. Engeser (University of Bonn), K. Kramer, A. Harzen, and H. Nakagami (MPI for Plant Breeding Research, Cologne), Susanne Neupert (Institute for Zoology, Cologne) and the Biomolecular Magnetic Resonance Spectroscopy facilities (University of Frankfurt) for technical support, training modules, and access to instruments. The FLI is a member of the Leibniz Association (WGL) and is financially supported by the Federal Government of Germany and the State of Thuringia. Financial support by a Liebig fellowship of the Fonds der Chemischen Industrie (Frankfurt, Germany) (to A.A.T.) and by the University of Bonn (to D.I.) is gratefully acknowledged. We also thank E. Krieger (YASARA Biosciences GmbH, Vienna, Austria) for useful scientific discussions.

References

- 1 J. S. Weissman and P. S. Kim, *Science*, 1991, **253**, 1386–1393.
- 2 T. E. Creighton, *Science*, 1992, **256**, 111–4.
- 3 J. S. Weissman and P. S. Kim, *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 9900–4.
- 4 J. Weissman and P. Kim, *Nat. Struct. Mol. Biol.*, 1995, **2**, 1123–30.
- 5 M. Cemazar, S. Zahariev, J. J. Lopez, O. Carugo, J. A. Jones, P. J. Hore and S. Pongor, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 5754–5759.
- 6 G. Bulaj, O. Buczek, I. Goodsell, E. C. Jimenez, J. Kranski, J. S. Nielsen, J. E. Garrett and B. M. Olivera, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 14562–14568.
- 7 A. I. Bartlett and S. E. Radford, *Nat. Struct. Mol. Biol.*, 2009, **16**, 582–588.
- 8 A. R. Fersht, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 650–654.
- 9 J. Y. Chang, *Biochemistry*, 2011, **50**, 3414–3431.
- 10 J. L. Dutton, P. S. Bansal, R. C. Hogg, D. J. Adams, P. F. Alewood and D. J. Craik, *J. Biol. Chem.*, 2002, **277**, 48849–48857.
- 11 A. A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Köhl, A. Mischo, G. Buntkowsky, M. Görlach, S. H. Heinemann and D. Imhof, *Angew. Chem. Int. Ed.*, 2012, **51**, 4058–4061.
- 12 B. M. Olivera, J. Rivier, J. K. Scott, D. R. Hillyard and L. J. Cruz, *J. Biol. Chem.*, 1991, **266**, 22067–22070.
- 13 B. B. Carstens, G. Berecki, J. T. Daniel, H. S. Lee, K. A. V. Jackson, H. S. Tae, M. Sadeghi, J. Castro, T. O'Donnell, A. Deiteren, S. M. Brierley, D. J. Craik, D. J. Adams and R. J. Clark, *Angew. Chem. Int. Ed.*, 2016, **55**, 4692–4696.
- 14 B. Leader, Q. J. Baca and D. E. Golan, *Nat. Rev. Drug. Discov.*, 2008, **7**, 21–39.

- 15 B. M. Olivera, J. Rivier, J. K. Scott, D. R. Hillyard and L. J. Cruz, *J. Biol. Chem.*, 1991, **266**, 22067–22070.
- 16 M. Góngora-Benítez, J. Tulla-Puche and F. Albericio, *Chem. Rev.*, 2014, **114**, 901–926.
- 17 S. Peigneur, M. Paolini-Bertrand, H. Gaertner, D. Biass, A. Violette, R. Stöcklin, P. Favreau, J. Tytgat and O. Hartley, *J. Biol. Chem.*, 2014, **289**, 35341–35350.
- 18 S. Dutertre, E. A. B. Undheim, S. S. Pineda, A.-H. Jin, V. Laverigne, B. G. Fry, R. J. Lewis, P. F. Alewood and G. F. King, in *Venoms to Drugs: Venom as a Source for the Development of Human Therapeutics*, ed. G. F. King, Wiley-VCH, Weinheim, 2015, pp. 80–96.
- 19 J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning and M. Mann, *Nat. Methods*, 2007, **4**, 709–12.
- 20 R. Aebersold and M. Mann, *Nature*, 2003, **422**, 198–207.
- 21 A. Szyk, Z. Wu, K. Tucker, D. Yang, W. Lu and J. Lubkowski, *Protein Sci.*, 2006, **15**, 2749–2760.
- 22 H. N. Hunter, D. B. Fulton, T. Ganz and H. J. Vogel, *J. Biol. Chem.*, 2002, **277**, 37597–37603.
- 23 T. Sitar, G. M. Popowicz, I. Siwanowicz, R. Huber and T. A. Holak, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 13028–13033.
- 24 T. Chang, F. Hsieh, M. Zebisch, K. Harlos, J. Elegheert and E. Y. Jones, *eLife*, 2015, 1–27.
- 25 H. W. Christinger, G. Fuh, A. M. de Vos and C. Wiesmann, *J. Biol. Chem.*, 2004, **279**, 10382–10388.
- 26 C. M. Steppan, S. T. Bailey, S. Bhat, E. J. Brown, R. R. Banerjee, C. M. Wright, H. R. Patel, R. S. Ahima and M. A. Lazar, *Nature*, 2001, **409**, 307–312.
- 27 J. L. Marshall, J. Holmberg, E. Chou, A. C. Ocampo, J. Oh, J. Lee, A. K. Peter, P. T. Martin and R. H. Crosbie-watson, *J. Cell Biol.*, 2012, **197**, 1009–1027.
- 28 M. S. Lyons, B. Bell, D. Stainier and K. G. Peters, *Dev. Dyn.*, 1998, **212**, 133–140.
- 29 A. Albert, J. J. Eksteen, J. Isaksson, M. Senge, T. Hansen and T. Vasskog, *Anal. Chem.*, 2016, **88**, 9539–9546.
- 30 I. Szabo, G. Schlosser, F. Hudecz and G. Mezo, *Biopolymers*, 2007, **88**, 20–28.
- 31 T.-Y. Yen, H. Yan and B. A. Macher, *J. Mass Spectrom.*, 2002, **37**, 15–30.
- 32 P. Heimer, A. A. Tietze, M. Böhm, R. Giernoth, A. Kuchenbuch, A. Stark, E. Leipold, S. H. Heinemann, C. Kandt and D. Imhof, *ChemBioChem*, 2014, **15**, 2754–2765.
- 33 K. J. Nielsen, M. Watson, D. J. Adams, A. K. Hammarstrom, P. W. Gage, J. M. Hill, D. J. Craik, L. Thomas, D. Adams, P. F. Alewood and R. J. Lewis, *J. Biol. Chem.*, 2002, **277**, 27247–27255.
- 34 L. Poppe, J. O. Hui, J. Ligutti, J. K. Murray and P. D. Schnier, *Anal. Chem.*, 2012, **84**, 262–266.
- 35 J. R. McArthur, V. Ostroumov, A. Al-sabi, D. McMaster and R. J. French, *Biochemistry*, 2011, **50**, 116–124.
- 36 R. Chen, A. Robinson and S. H. Chung, *PLoS One*, 2014, **9**, 1–8.
- 37 J. Wu, Z. Yan, Z. Li, C. Yan, S. Lu, M. Dong and N. Yan, *Nature*, 2016, **537**, 191–196.
- 38 H. Shen, H. Shen, Q. Zhou, X. Pan, Z. Li, J. Wu and N. Yan, *Science*, 2017, **4326**, 1–18.
- 39 C. Dominguez, R. Boelens and A. Bonvin, *J. Am. Chem. Soc.*, 2003, **125**, 1731–1737.
- 40 S. J. de Vries, A. D. J. van Dijk, M. Krzeminski, M. van Dijk, A. Thureau, V. Hsu, T. Wassenaar and A. M. J. J. Bonvin, *Proteins*, 2008, **69**, 726–733.
- 41 J. R. McArthur, G. Singh, M. L. O'Mara, D. McMaster, V. Ostroumov, D. P. Tieleman and R. J. French, *Mol. Pharmacol.*, 2011, **80**, 219–227.
- 42 S. Mahdavi and S. Kuyucak, *Toxins*, 2014, **6**, 3454–3470.
- 43 S. Mahdavi and S. Kuyucak, *PLoS One*, 2014, **9**, 1–12.
- 44 P. Heimer, A. A. Tietze, M. Böhm, R. Giernoth, A. Kuchenbuch, A. Stark, E. Leipold, S. H. Heinemann, C. Kandt and D. Imhof, *ChemBioChem*, 2014, **15**, 2754–2765.
- 45 W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, *Proteins*, 2005, **59**, 687–696.
- 46 M. S. Cheung, M. L. Maguire, T. J. Stevens and R. W. Broadhurst, *J. Magn. Reson.*, 2010, **202**, 223–233.
- 47 E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein and G. Vriend, *Proteins Struct. Funct. Bioinf.*, 2004, **57**, 678–683.
- 48 E. Krieger, K. Joo, J. Lee, J. Lee, S. Raman, J. Thompson, M. Tyka, D. Baker and K. Karplus, *Proteins Struct. Funct. Bioinf.*, 2009, **77**, 114–122.
- 49 E. Krieger, G. Koraimann and G. Vriend, *Proteins Struct. Funct. Bioinf.*, 2002, **47**, 393–402.
- 50 E. Krieger, J. E. Nielsen, C. A. Spronk and G. Vriend, *J. Mol. Graph. Model.*, 2006, **25**, 481–486.
- 51 F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson and D. G. Higgins, *Mol. Syst. Biol.*, 2011, **7**, 1–6.
- 52 J. Payandeh, T. Scheuer, Z. N. and C. W. a, *Nature*, 2011, **475**, 353–358.
- 53 A. S. Konagurthu, J. C. Whisstock, P. J. Stuckey and A. M. Lesk, *Proteins*, 2006, **64**, 559–574.
- 54 A. A. Canutescu, A. A. Shelenkov and R. L. Dunbrack Jr, *Protein Sci.*, 2003, **12**, 2001–2014.
- 55 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 56 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712–725.
- 57 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- 58 S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952–962.
- 59 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
- 60 E. Krieger and G. Vriend, *J. Comput. Chem.*, 2015, **36**, 996–1007.
- 61 R. Zhao, H. Dai, S. Qiu, T. Li, Y. He, Y. Ma, Z. Chen and Y. Wu, *PLoS One*, 2011, **6**, 1–10.

Epilogue

The 15 possible μ -PIIIA isomers possessing each three disulfide bonds have been synthesized and characterized for the first time. Comparison of the analytical methods subjected revealed that HPLC and NMR alone are not sufficient to distinguish between all isomers indicating the need for better purification methods or alternative procedures. In contrast, MALDI-MS/MS was the only method to identify the respective disulfide connectivity, yet this was highly depending on the purity of the peptides after partial reduction and derivatization. In case of μ -PIIIA, LC-ESI-MS/MS was not suitable for full fragmentation of the linear peptide sequence contrary to the earlier described conotoxins μ -SIIIA and δ -EVIA. It was further found that the biological activity of the isomers seemed to correlate with the rigidity of the synthesized structure with the native connectivity as the most active one within a set of further active peptide isomers. More than 50% of the isomers investigated showed no significant block of the ion channel Na_v1.4.

This article improves the knowledge regarding μ -conotoxins by focusing on the impact of the disulfide bonds on structure elucidation and biological activity in contrast to the previous publication which concentrated on sequence dependence. Furthermore, a direct comparison of different analytical methods was carried out in order to answer the question if one method alone is sufficient to identify all 15 possible isomers of a μ -conotoxin. As a consequence of this study the urgent need for suitable standards of multiple disulfide-bridged peptides is derived.

5. Conclusions

This work comprises two parts. The first chapter summarizes publications of parts I, II and III reporting the results from studies of the oxidative folding of cysteine-rich peptides in room-temperature Ionic Liquids. The second chapter (parts III – VI) combines structure elucidation and analytical characterization with respective structure-activity relationship studies of selected conotoxins.

5.1 Oxidative folding of cysteine-rich peptides in Ionic Liquids

Ionic Liquids were established as novel solvents for various applications including peptide chemistry, e.g. in the synthesis and characterization of peptides.¹⁶⁹ In this work oxidative folding of cysteine-containing conopeptides in Ionic Liquids was in the focus and was compared with conventional buffer oxidation using redox-active agents. Several features were identified, which influenced the product formation during the oxidation process. On the one hand, the primary sequence of the peptide including its length and net charge and, on the other hand, the RT-IL used for the oxidation had the largest influence on product formation depending on the hydrophilicity and hydrophobicity, respectively, of the peptide. For the conotoxins of interest, imidazolium-based RT-ILs showed best results. However, the anion had a large impact on the oxidation outcome. While acetate resulted in high oxidation yields, diethylphosphate, dicyanamide and tosylate did not lead to product formation for various conotoxins. The impact of IL ions on solvation of peptides and proteins have been described earlier.²³⁸ In addition, the water content of the respective RT-IL (>3% v/v) decreased oxidation efficacy of the applied peptides, while for proteins or enzymes higher concentrations of water are required to maintain the correct folding for bioactivity or catalytic function.¹⁶⁹ Higher temperatures (60 – 80 °C) increased the oxidation rate and prevented the peptide from thermal decomposition resulting in a faster product formation compared to room temperature reactions. Although much higher peptide concentrations were applied for IL-oxidation, the intramolecular formation of disulfide bonds was preferred in contrast to intermolecular formation of di- or oligomers due to the unique solvation effects of IL ions. The block of the ion channel Na_v1.4 by IL-oxidized μ -conotoxin SIIIA could be demonstrated, however, the biological activity was drastically reduced. This effect was explained by IL ions attached to amino acid side chains responsible for the interaction between toxin and ion channel. This hypothesis was supported by MD simulations showing acetate ions interacting

with basic residues, e.g. Lys and Arg, of the sequence responsible for bioactivity. Attachment of IL ions to peptides influencing the structure was also reported by Lesch *et al.*²³⁹ Here, it was observed by a shift in the retention time after multiple purification steps in HPLC analysis indicating that the ions remained on the peptide. This shift was also observed for the peptide oxidized in buffer and stored in an IL, which also revealed a reduced bioactivity. Furthermore, formation of different disulfide-bridged isomers could be excluded after determination of the disulfide connectivity via LC-ESI-MS/MS, which is discussed in more detail in the chapter 5.2. Therefore, the preparation of peptides and proteins in RT-ILs show significant limitations with respect to their biological function or target activity. However, although the inhibitory effect of the toxin μ -SIIIA on the ion channel was hindered by the attached IL ions, the peptide showed no limitations with respect to a tryptic digest using BAPNA as the chromogenic substrate. Here, access of the enzyme to the peptide was not hindered by the IL ions attached.

Although ILs have high potential for several applications, their application for oxidative folding of conotoxins is limited because the product obtained has only decreased bioactivity compared to conventional, buffer-oxidized peptide even after several purification steps. A solution for this limitation would be an improved purification method to remove the attached ions and release of the fully bioactive compound.

5.2 Analytical characterization and structure-activity relationships of conotoxins

Detailed analytical characterizations in combination with structure elucidation are required to confirm the synthesized conotoxin structure and are essential for further SAR studies with the respective biological target. One important aspect is the analytical confirmation of the disulfide connectivity in the conotoxins of interest, and cysteine-containing peptides and proteins in general. Therefore, a partial reduction and derivatization protocol in combination with LC-ESI-MS/MS was applied to μ -SIIIA oxidized in buffer solution, oxidized in IL and buffer-oxidized peptide stored in IL to confirm the formation of the native disulfide connectivity. Partial reduction was performed with TCEP and alkylation of the thiol-function with IAA revealed the former disulfide bridging. In all cases the same disulfide crosslinking was observed for the tested peptides. In addition, small amounts of other isomers could be identified using this method. In addition, identification of the specific conotoxin binding sites were rarely investigated so far, because of their exhausting and time-consuming manner. However, such studies are pivotal for clarifying the respective target selectivity and preventing unwanted side effects resulting from unspecific binding to other targets.

5. Conclusions

The δ -conotoxin EVIA selectively inhibits the inactivation of neuronal Na⁺-channel Na_v1.7 without effecting muscular or cardiac muscular VGSCs (Na_v1.4 and 1.5). This enhances terminal excitability of nerves and synaptic efficacy at the neuromuscular junction by increasing the duration of action potentials in myelinated axons and spinal neurons.²⁴⁰ Although the structure was published previously by Volpon *et al.*²¹⁰ the structure of the self-synthesized δ -EVIA was confirmed by NMR studies. Determination of the disulfide connectivity was performed with the same protocol as previously described for μ -SIIIA using partial reduction and derivatization prior to sequencing via LC-ESI-MS/MS. The toxin binding site explaining the selectivity of the toxin was identified by a combination of theoretical molecular modelling and docking studies and experimental patch-clamp experiments using chimeras of sensitive Na_v1.7 and insensitive Na_v1.4. This combined approach identified domain I and IV as pivotal for the sensitivity of Na_v1.7. In addition, the linker between segment 3 and 4 was identified as most important part for bioactivity within domain IV because of the interaction between inactivated voltage sensor and δ -EVIA decreasing speed of inactivation. These results improve the knowledge of voltage-dependent channel gating and the mechanism of activation and inactivation for respective isoforms. However, further research is required to fully understand the complete mechanisms of activation and inactivation of VGSCs. Here, the δ -conotoxins represent an interesting research tool for further investigations, but also to support studies of other voltage-sensor toxins obtained from scorpions, spiders or sea anemones.

Although intensive studies of the past focused on μ -conotoxins, the reason for their high subtype specificity within VGSCs, which is conserved in a short sequence (15 - 25 AA), is not fully understood yet. Furthermore, the product formation during self-folding seems to be highly sequence-dependent resulting in one or more disulfide-bridged isomers.^{63,143} Besides the native conformation other isomers obtained for μ -KIIIA and μ -PIIIA revealed biological activity too, complicating the whole process of establishing reliable SARs. To study the impact of the amino acid sequence and formation of disulfide bonds on the formation of the folded peptide a combination of *in silico* and *in vitro* experiments was used. Therefore, five μ -conotoxins (GIIIA, KIIIA, PIIIA, SIIIA and SmIIIA) were used as model peptides. The product formation during oxidation with redox-active agents described in previous studies was confirmed indicating two groups.^{63,91,116} One major product was observed for μ -GIIIA and μ -SIIIA harbouring the native connectivity, whereas μ -PIIIA and μ -SmIIIA resulted in several fully oxidized isomers. For the remaining μ -KIIIA one major product with a non-native crosslinking (Cys1-Cys5, Cys2-Cys4, Cys3-Cys6) besides one minor product containing the native bridging were observed. The formation of different products can be explained with two existing folding mechanisms. On the one hand, a rapid collapse provokes fast formation of a

stable conformation, while on the other hand a slower re-arrangement results in the formation of different isomers for the individual conotoxins. Molecular dynamics simulations were utilized to receive an impression about this process. Starting from the respective NMR structures the removal of single disulfide bonds elucidated the conformational stability of the structures. Distance measurements of the free thiol functions for the one and two disulfide linked peptides in combination with RMSD and RMSF reveal the rigidity of the structure. From this research the contribution of the primary sequence to the final stability is ascertained. Furthermore, the rigidity of the disulfide-deficient species can be used to estimate their biological activity especially for the one bridge opened version. Here, the disulfide bond Cys2-Cys5 had no significant effect with respect to flexibility of the structure. The remaining two bridges (Cys1-Cys4, Cys3-Cys6) between the adjacent cysteines were sufficient to maintain the conformational stability. This implemented that this bond can be removed without significant or complete loss of biological activity. Although extensive research studies are missing to test all five conotoxins, the results obtained from previous studies with μ -KIIIA support our findings with respect to conformational stability and biological activity.^{241,242} Recent developments for disulfide-deficient species of μ -GIIIA revealed the influence of the individual disulfide bridges upon alanine mutation.²⁴³ After removing the Cys2-Cys5 connectivity biological activity on Na_v1.4 was still observed. However, a reduced bioactivity (e.g. up to 50%) was observed for all mutations leaving open the impact of alanine substitutions on the rigidity of the structure. i.e. serine would probably represent a more suitable residue for future mutational studies.

To clarify the impact of the residues on conformational stability of the conotoxins a visualization approach was applied in order to change the scope to more global residue movement observations. Therefore, residue surface proximity studies rendering the native structure in layers for the whole MD simulation time were performed. The layer membership of individual residues changes during simulation time influencing the adjacent residues. Here, the basic residues, essential for biological activity, maintain their position on the surface of the molecule. Instead the cysteine residues, responsible for structural stability, and also serine changes layer membership rather often. This opens a new perspective, revealing a paradox view on the commonly stable disulfide bonds moving frequently to stabilize the orientation of the basic residues on the surface of the molecule. Although such a layer approach was utilized for larger proteins before,²⁴⁴ the application to peptides reveals new insights into the movement of individual residues. Finally, the complex folding behaviour of disulfide-rich peptides holds great potential for future research to elucidate mechanisms of product (correctly folded) and by-product (misfolded) formation as well as simplification of the corresponding conotoxins to simpler synthetic versions. Research within the field of *in silico* studies help to understand and show complex processes, e.g. folding, docking and dynamic

5. Conclusions

simulations, to a broader scientific community. However, the results of such studies have to be verified with experimental data to ensure the correctness of the obtained data.

After analyzing the influence of the primary sequence on the biological activity of μ -conotoxins, the residues are kept constant to elucidate only the impact of a different disulfide crosslinking on the corresponding activity. In previous studies by Tietze *et al.*, three isomers of μ -PIIIA were isolated from a self-folding approach and characterized.⁹⁶ Besides the native-bridged isomer another isomer harbouring the disulfide connectivity Cys1-Cys5, Cys2-Cys6, Cys3-Cys4, which was confirmed by 2D-NMR experiments, showed a higher biological activity on Na_v1.4 compared to the native one. Therefore, μ -PIIIA was chosen as a model peptide to proceed with the further work in this direction.

It could be shown that one method alone is not suitable for structure analysis of μ -PIIIA isomers, and, in turn, for venomics approaches because each method used has a unique output, i.e. confirmation of primary sequence, disulfide connectivity or 3D conformation lacking the combination of all aspects to distinguish between disulfide conformational isomers. In some cases, these limitations of the respective methods clearly depend on the primary sequence examined, which might result in ambiguous or even misleading information as was described for several examples, e.g. the disulfide connectivity of μ -KIIIA.¹¹⁶

The evaluation of all possible 15 μ -PIIIA isomers is even more time consuming compared to three isomers derived from an α -conotoxin, because in the latter case all isoforms can possibly be baseline-separated in HPLC analysis. Thus, single forms can be isolated and characterized. This is indeed not the case for the disulfide-rich peptides containing three disulfide bonds, which was demonstrated in this work by co-elution of different isomers showing similar or even identical elution behaviour. In some cases the characterization of one single peak lead to ambiguous results due to contamination with other isomers. Nevertheless, for the first time 15 isomers of a 3-times disulfide-linked conotoxin were synthesized, chemically and structurally analyzed and tested for their activity at skeletal Na_v1.4. Discrimination of the different isomers with MALDI-MS/MS was possible depending on the purity of the sample, whereas ESI-MS/MS was inefficient for the sequencing of μ -PIIIA because of low sequence coverage. It was shown that the potential of the available analytical methods (HPLC, NMR, MALDI-MS/MS and ESI-MS/MS) is limited and an individual method is not able to discriminate unequivocally between the 15 isomers. Thus, this study demonstrates the urgent need for suitable standards of peptides and proteins with multiple disulfide bonds to verify a disulfide connectivity obtained experimentally.

In the further process of SAR studies the obtained μ -PIIIA structures were clustered in three groups according to their structural flexibility. In docking studies the rigid isomers revealed a higher capacity of binding to Nav1.4 ion channel compared to the flexible structures. These results correlated well with the observations from the electrophysiological experiments, in which the rigid isomers Cys1-Cys5, Cys2-Cys3, Cys4-Cys6 and the native connectivity Cys1-Cys4, Cys2-Cys5, Cys3-Cys6 were most active, while the most flexible isomers Cys1-Cys2, Cys3-Cys4, Cys5-Cys6 and Cys1-Cys3, Cys2-Cys6, Cys4-Cys5 were inactive. Furthermore, the resting time of the different structures on the channel pore also correlated with the experimentally determined time constants. These observations were attributed to the different orientations of the basic residues Arg2, Arg12, Arg14, Arg20, Lys9 and Lys17, as well as Phe7 and Pyr1 and their interaction with the surface of the channel and coverage of the pore region. Especially Arg14, which interacts with the selectivity filter motif, and Arg20, which protrudes the pore, were essential for high binding affinity.

In conclusion, this thesis comprises studies covering multiple aspects of conotoxin research including synthesis, oxidation, structure elucidation, SAR studies and biological activity on VGSCs. In addition, the advantages and limitations of a set of analytical methods with respect to the determination of disulfide bridging of peptides was critically evaluated.

6. Abbreviations

All Abbreviations of amino acids and their derivatives are according to the recommendation of the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) and the Nomenclature Committee of IUB (NC-IUB).²⁴⁵ In general, *L*-configuration of amino acids and amino acid derivatives was used.

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
1,5-DAN	1,5- diaminonaphthalene
ACE	angiotensin-converting enzyme
BAPNA	(N α -benzoyl- <i>DL</i> -arginine-4-nitroanilide hydrochloride)
Boc	<i>tert</i> -butyloxycarbonyl
BPTI	bovine pancreatic trypsin inhibitor
[C ₂ mim][OAc]	1-ethyl-3-methylimidazolium acetate
CID	collision induced decay
Cys or C	cysteine
COSY	correlated spectroscopy
CRISP	cysteine-rich secretory protein
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
ESI	electrospray ionization
FDA	food and drug administration
Fmoc	9-fluorenylmethyloxycarbonyl
GABA	γ -Aminobutyric acid
GLP-1	glucagon-like peptide-1
GSH	glutathione reduced
GSSG	glutathione oxidized
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence
ICK	inhibitory cysteine knot
IL	Ionic Liquid
ISD	in source decay

LC	liquid chromatography
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
MS/MS	tandem mass spectrometry
nAChR	nicotinic acetylcholine receptor
NCL	native chemical ligation
NEM	N-ethylmaleimide
NMR	nuclear magnetic resonance
NOE	nuclear overhauser enhancement
NOESY	nuclear overhauser and enhancement spectroscopy
OS	organic solvent
PDB	protein data bank
PDI	protein disulfide isomerases
PG	protecting group
RMSD	root-mean-square-deviation
ROE	rotating-frame overhauser enhancement
ROESY	rotating-frame nuclear overhauser enhancement spectroscopy
RP	reversed-phase
RT-IL	room-temperature Ionic Liquid
SAR	structure-activity relationship
SPS	solution-phase peptide synthesis
SPPS	solid-phase peptide synthesis
TCEP	<i>tris</i> (2-carboxyethyl)phosphine
TFA	trifluoro acetic acid
TOCSY	total correlated spectroscopy
TTX	tetrodotoxin
VGSC	voltage-gated sodium channels

7. List of figures

Figure 1. Advantage of constrained peptides as therapeutics, filling the gap between small molecules and biologics by combining the properties of each type of drugs. Modified from ¹. 3

Figure 2. Shells from different fish-hunting cone snails: *Conus ermineus* (left), *Conus purpurascens* (middle) and *Conus striatus* (right). Modified from ^{68,69} 5

Figure 3. (Left) Schematic view of a cone snail venomous apparatus. (Right) Harpoon-like radular teeth enlarged by scanning electron micrograph from *Conus purpurascens*. Modified from ⁴⁵ 6

Figure 4. Crude venom of *Conus geographus* purified with HPLC and fractions labelled with respective biological activity. This was observed by intracranially injection of ~0.5 – 2 nmol of each fraction into mice. The fractions with † caused death at in least one injected animal. Modified from ⁴⁹ 7

Figure 5. Conotoxin classification displaying the relationship between genetic superfamily (red), pharmacological family (blue), cysteine framework (black) and cysteines included with respective connectivity if known (green). Modified from ^{56–58} 9

Figure 6. Voltage gated sodium channel α -subunit consisting of 4 homologous domains (D1-4) with each consisting of 6 transmembrane segments (S1-6). The voltage sensor S4 is marked in red and the channel forming units S5/6 are shown in green. The loop responsible for fast inactivation between D3 and D4 is labelled with the relevant residues IFM relevant for fast inactivation. Modified from ⁸⁶12

Figure 7. Topology of the three isolated μ -PIIIA isomers from oxidative folding and structures derived from NMR elucidation (A) with the respective disulfide connectivity (B). The individual blocking of VGSC Nav1.4 is described below (C) with the highest bioactivity for a non-native isoform. Modified from ⁹⁶16

Figure 8. Analytical methods for the elucidation of disulfide bond connectivity of small, disulfide-rich peptides and respective characterization opportunities. Samples subjected to X-ray crystallography or NMR can be characterized by destructive methods afterwards because recovering is possible, which is not possible vice versa. Modified from ¹21

8. List of tables

Table 1. Subtype specificity of different α -conotoxins on VGSCs relating to their IC₅₀ values.
Only KD available. CNS = central nervous system, DRG = dorsal root ganglia, PNS = peripheral nervous system; VGSC from *R. norvegicus* was used in general unless otherwise mentioned; * indicates bioactivity for respective isomer isolated; **1.6 from *M. musculus*. ...13

Table 2. Comparison of disulfide elucidation methods and sample preparation techniques applied to respective conotoxins.....25

9. References

- (1) Góngora-Benítez, M., Tulla-Puche, J., Albericio, F. Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chem. Rev.* **2014**, *114*, 901-926.
- (2) Goodwin, D., Simerska, P., Toth, I. Peptides as therapeutics with enhanced bioactivity. *Curr. Med. Chem.* **2012**, *19*, 4451-4461.
- (3) Leader, B., Baca, Q.J., Golan, D.E. Protein therapeutics: A summary and pharmacological classification. *Nat. Rev. Drug Discov.* **2008**, *7*, 21-39.
- (4) Escoubas, P., King, G.F. Venomics as a drug discovery platform. *Expert Rev. Proteomics* **2009**, *6*, 221-224.
- (5) Akondi, K.B., Muttenthaler, M., Kaas, Q., Craik, D.J., Lewis, R.J., Alewood, P.F. Discovery, synthesis, and structure-activity relationships of conotoxins. *Chem. Rev.* **2014**, *114*, 5815-5847.
- (6) Olivera, B.M., Gray, W.R., Zeikus, R., McIntosh, J.M., Varga, J., Rivier, J., de Santos, V., Cruz, L.J. Peptide neurotoxins from fish-hunting cone snails. *Science* **1985**, *230*, 1338-1343.
- (7) Sanford, M. Intrathecal Ziconotide: A review of its use in patients with chronic pain refractory to other systemic or intrathecal analgesics. *CNS Drugs* **2013**, *27*, 989-1002.
- (8) Gorson, J., Ramrattan, G., Verdes, A., Wright, E.M., Kantor, Y., Rajaram Srinivasan, R., Musunuri, R., Packer, D., Albano, G., Qiu, W.G., Holford, M. Molecular diversity and gene evolution of the venom arsenal of Terebridae predatory marine snails. *Genome Biol. Evol.* **2015**, *7*, 1761-1778.
- (9) Becker, S., Terlau, H. Toxins from cone snails: Properties, applications and biotechnological production. *Appl. Microbiol. Biotechnol.* **2008**, *79*, 1-9.
- (10) Kang, T.S., Kini, R.M. Structural determinants of protein folding. *Cell. Mol. Life Sci.* **2009**, *66*, 2341-2361.
- (11) Paracelsus, T. Septem defensiones. **1965**, 508-513.
- (12) Edwards, A.M., Bountra, C., Kerr, D.J., Willson, T.M. Open access chemical and clinical probes to support drug discovery. *Nat. Chem. Biol.* **2009**, *5*, 436-440.
- (13) Rodrigues, T., Reker, D., Schneider, P., Schneider, G. Counting on natural products for drug design. *Nat. Chem.* **2016**, *8*, 531-541.
- (14) Vetter, I., Davis, J.L., Rash, L.D., Anangi, R., Mobli, M., Alewood, P.F., Lewis, R.J., King, G.F. Venomics: A new paradigm for natural products-based drug discovery. *Amino Acids* **2011**, *40*, 15-28.
- (15) Fry, B.G., Roelants, K., Champagne, D.E., Scheib, H., Tyndall, J.D., King, G.F., Nevalainen, T.J., Norman, J., Lewis, R.J., Norton, R.S., Renjifo, C., de la Vega, R.C.R. The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 483-511.
- (16) Lewis, R.J., Garcia, M.L. Therapeutic potential of venom peptides. *Nat. Rev. Drug Discov.* **2003**, *2*, 790-802.
- (17) Harvey, A.L. Toxins and drug discovery. *Toxicon* **2014**, *92*, 193-200.
- (18) Hughes, J.P., Rees, S.S., Kalindjian, S.B., Philpott, K.L. Principles of early drug discovery. *Br. J. Pharmacol.* **2011**, *162*, 1239-1249.
- (19) Favreau, P., Stöcklin, R. Marine snail venoms: Use and trends in receptor and channel neuropharmacology. *Curr. Opin. Pharmacol.* **2009**, *9*, 594-601.
- (20) Cushman, D.W., Ondetti, M.A. History of the design of Captopril and related inhibitors of angiotensin converting enzyme. *Hypertension* **1991**, *17*, 589-592.
- (21) Schroeder, C.I., Craik, D.J. Therapeutic potential of conopeptides. *Future Med. Chem.* **2012**, *4*, 1243-1255.
- (22) Colgrave, M.L., Craik, D.J. Thermal, chemical, and enzymatic stability of the cyclotide Kalata B1: The

- p>importance of the cyclic cystine knot.
- Biochemistry*
- 2004**
- ,
- 43*
- , 5965-5975.
- (23) Craik, D.J., Daly, N.L., Waine, C. The cystine knot motif in toxins and implications for drug design. *Toxicon* **2001**, *39*, 43-60.
 - (24) Dardevet, L., Rani, D., El Aziz, T.A., Bazin, I., Sabatier, J.M., Fadl, M., Brambilla, E., De Waard, M. Chlorotoxin: A helpful natural scorpion peptide to diagnose glioma and fight tumor invasion. *Toxins* **2015**, *7*, 1079-1101.
 - (25) Malone, J., Trautmann, M., Wilhelm, K., Taylor, K., Kendall, D.M. Exenatide once weekly for the treatment of type 2 diabetes. *Expert Opin. Investig. Drugs* **2009**, *18*, 359-367.
 - (26) Du Vigneaud, V., Ressler, C., Swan, J., Roberts, C., Katsoyannis, P., Gordon, S. The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* **1953**, *75*, 4879-4880.
 - (27) Merrifield, R.B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
 - (28) Isidro-Llobet, A., Alvarez, M., Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **2009**, *109*, 2455-2504.
 - (29) Guzmán, F., Barberis, S., Illanes, A. Peptide synthesis: Chemical or enzymatic. *Electron. J. Biotechnol.* **2007**, *10*, 279-314.
 - (30) Muir, T.W., Dawson, P.E., Kent, S.B. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzym.* **1997**, *289*, 266-298.
 - (31) Wieland, T., Bokelmann, E., Bauer, L., Lang, H.U., Lau, H. Über Peptidsynthesen. 8. Mitteilung Bildung von S-haltigen Peptiden durch intramolekulare Wanderung von Aminoacylresten. *Liebigs Ann. Chem.* **1953**, *583*, 129-149.
 - (32) Dawson, P.E., Muir, T.W., Clark-Lewis, I., Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776-779.
 - (33) Dawson, P.E., Kent, S.B. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923-960.
 - (34) Tam, J.P., Yu, Q., Miao, Z. Orthogonal ligation strategies for peptide and protein. *Biopolymers* **1999**, *51*, 311-332.
 - (35) Moroder, L., Musiol, H.J., Götz, M., Renner, C. Synthesis of single- and multiple-stranded cystine-rich peptides. *Biopolymers* **2005**, *80*, 85-97.
 - (36) Kohn, A.J. Piscivorous gastropods of the genus *Conus*. *Proc. Natl. Acad. Sci. USA* **1956**, *42*, 168-171.
 - (37) Rice, R.D., Halstead, B.W. Report of fatal cone shell sting by *Conus geographus linneaus*. *Toxicon* **1968**, *5*, 223-224.
 - (38) Terlau, H., Olivera, B.M. *Conus* venoms: A rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **2004**, *84*, 41-68.
 - (39) Phuong, M.A., Mahardika, G.N., Alfaro, M.E. Dietary breadth is positively correlated with venom complexity in cone snails. *BMC Genomics* **2016**, *17*, 401.
 - (40) Kohn, A.J. Maximal species richness in *Conus*: Diversity, diet and habitat on reefs of northeast Papua New Guinea. *Coral Reefs* **2001**, *20*, 25-38.
 - (41) Puillandre, N., Duda, T.F., Meyer, C., Olivera, B.M., Bouchet, P. One, four or 100 genera? A new classification of the cone snails. *J. Molluscan Stud.* **2015**, *81*, 1-23.
 - (42) Cruz, L.J., Gray, W.R., Yoshikami, D., Olivera, B.M. *Conus* venoms: A rich source of neuroactive peptides. *J. Toxicol. Toxin Rev.* **1985**, *4*, 107-132.
 - (43) Dutertre, S., Jin, A.H., Alewood, P.F., Lewis, R.J. Intraspecific variations in *Conus geographus* defence-evoked venom and estimation of the human lethal dose. *Toxicon* **2014**, *91*, 135-144.
 - (44) Lavergne, V., Harliwong, I., Jones, A., Miller, D., Taft, R.J., Alewood, P.F. Optimized deep-targeted proteo-transcriptomic profiling reveals unexplored *Conus* toxin diversity and novel cysteine frameworks. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3782-E3791.

9. References

- (45) Olivera, B.M. *Conus* venom peptides: Reflections from the biology of clades and species. *Annu. Rev. Ecol. Syst.* **2002**, *33*, 25-47.
- (46) Duda, T.F., Chang, D., Lewis, B.D., Lee, T. Geographic variation in venom allelic composition and diets of the widespread predatory marine gastropod *Conus ebraeus*. *PLoS One* **2009**, *4*, 1-8.
- (47) Neves, J., Campos, A., Osorio, H., Antunes, A., Vasconcelos, V. Conopeptides from Cape Verde *Conus crotchii*. *Mar. Drugs* **2013**, *11*, 2203-2215.
- (48) Rodriguez, A.M., Dutertre, S., Lewis, R.J., Marí, F. Intraspecific variations in *Conus purpurascens* injected venom using LC/MALDI-TOF-MS and LC-ESI-TripleTOF-MS. *Anal. Bioanal. Chem.* **2015**, *407*, 6105-6116.
- (49) Olivera, B.M., Rivier, J., Clark, C., Ramilo, C.A., Corpuz, G.P., Abogadie, F., Mena, E., Woodward, S., Hillyard, D.R., Cruz, L.J. Diversity of *Conus* neuropeptides. *Science* **1990**, 257-263.
- (50) Biass, D., Violette, A., Hulo, N., Lisacek, F., Favreau, P., Stöcklin, R. Uncovering intense protein diversification in a cone snail venom gland using an integrative venomomics approach. *J. Proteome Res.* **2015**, *14*, 628-638.
- (51) Spence, I., Gillissen, D., Gregson, R.P., Quinn, R.J. Characterization of the neurotoxic constituents of *Conus geographus* (I) venom. *Life Sci.* **1977**, *21*, 1759-1769.
- (52) Olivera, B.M., Cruz, L.J., Lecheminant, G.W., Griffin, D., Zeikus, R., McIntosh, J.M., Galyean, R., Varga, J., Gray, W.R., Rivier, J. Neuronal calcium channel antagonists. Discrimination between calcium channel subtypes using ω -conotoxin from *Conus magus* venom. *Biochemistry* **1987**, *26*, 2086-2090.
- (53) Clark, C., Olivera, B.M., Cruz, L.J. A toxin from the venom of the marine snail *Conus geographus* which acts on the vertebrate central nervous system. *Toxicon* **1981**, *19*, 691-699.
- (54) Puillandre, N., Koua, D., Favreau, P., Olivera, B.M., Stöcklin, R. Molecular phylogeny, classification and evolution of conopeptides. *J. Mol. Evol.* **2012**, *74*, 297-309.
- (55) Woodward, S.R., Cruz, L.J., Olivera, B.M., Hillyard, D.R. Constant and hypervariable regions in conotoxin propeptides. *EMBO J.* **1990**, *9*, 1015-1020.
- (56) Robinson, S.D., Norton, R.S. Conotoxin gene superfamilies. *Mar. Drugs* **2014**, *12*, 6058-6101.
- (57) Kaas, Q., Westermann, J.C., Craik, D.J. Conopeptide characterization and classifications: An analysis using ConoServer. *Toxicon* **2010**, *55*, 1491-1509.
- (58) Wu, X., Huang, Y.H., Kaas, Q., Craik, D.J. Cyclisation of disulfide-rich conotoxins in drug design applications. *Eur. J. Org. Chem.* **2016**, *21*, 3462-3472.
- (59) Gray, W.R., Luque, A., Olivera, B.M., Barrett, J., Cruz, L.J. Peptide toxins from *Conus geographus* venom. *J. Biol. Chem.* **1981**, *256*, 4734-4740.
- (60) Azam, L., McIntosh, J.M. α -Conotoxins as pharmacological probes of nicotinic acetylcholine receptors. *Acta Pharmacol. Sin.* **2009**, *30*, 771-783.
- (61) McIntosh, J.M., Santos, A.D., Olivera, B.M. *Conus* peptide targeted to specific nicotinic acetylcholine receptor subtypes. *Annu. Rev. Biochem.* **1999**, *68*, 59-88.
- (62) Olivera, B.M., Rivier, J., Scott, J.K., Hillyard, D.R., Cruz, L.J. Conotoxins. *J. Biol. Chem.* **1991**, *266*, 22067-22070.
- (63) Fuller, E., Green, B.R., Catlin, P., Buczek, O., Nielsen, J.S., Olivera, B.M., Bulaj, G. Oxidative folding of conotoxins sharing an identical disulfide bridging framework. *FEBS J.* **2005**, *272*, 1727-1738.
- (64) Nielsen, K.J., Watson, M., Adams, D.J., Hammarström, A.K., Gage, P.W., Hill, J.M., Craik, D.J., Thomas, L., Adams, D., Alewood, P.F., Lewis, R.J. Solution structure of μ -conotoxin PIIIA, a preferential inhibitor of persistent tetrodotoxin-sensitive sodium channels. *J. Biol. Chem.* **2002**, *277*, 27247-27255.
- (65) Corpuz, G.P., Jacobsen, R.B., Jimenez, E.C., Watkins, M., Walker, C., Colledge, C., Garrett, J.E., McDougal, O., Li, W., Gray, W.R., Hillyard, D.R., Rivier, J., McIntosh, J.M., Cruz, L.J., Olivera, B.M. Definition of the M-conotoxin superfamily: Characterization of novel peptides from molluscivorous *Conus* venoms. *Biochemistry* **2005**, *44*, 8176-8186.

- (66) Jacob, R.B., McDougal, O.M. The M-superfamily of conotoxins: A review. *Cell. Mol. Life Sci.* **2010**, *67*, 17-27.
- (67) de la Cruz, R., Whitby, F.G., Buczek, O., Bulaj, G. Detergent-assisted oxidative folding of δ -conotoxins. *J. Pept. Res.* **2002**, *61*, 202-212.
- (68) Kaas, Q., Westermann, J.C., Halai, R., Wang, C.K.L., Craik, D.J. ConoServer, a database for conopeptide sequences and structures. *Bioinformatics* **2008**, *24*, 445-446.
- (69) Kaas, Q., Yu, R., Jin, A.H., Dutertre, S., Craik, D.J. ConoServer: Updated content, knowledge, and discovery tools in the conopeptide database. *Nucleic Acids Res.* **2012**, *40*, 325-330.
- (70) Endean, R., Gyr, P., Surridge, J. The pharmacological actions on guinea-pig ileum of crude venoms from the marine gastropods *Conus striatus* and *Conus magus*. *Toxicon* **1977**, *15*, 327-337.
- (71) Endean, R., Surridge, J., Gyr, P. Some effects of crude venom from the cones *Conus striatus* and *Conus magus* on isolated guinea-pig atria. *Toxicon* **1977**, *15*, 369-374.
- (72) Vetter, I., Lewis, R.J. Therapeutic potential of cone snail venom peptides (Conopeptides). *Curr. Top. Med. Chem.* **2012**, *12*, 1546-1552.
- (73) Wang, F., Yan, Z., Liu, Z., Wang, S., Wu, Q., Yu, S., Ding, J., Dai, Q. Neuropharmacology molecular basis of toxicity of N-type calcium channel inhibitor MVIIA. *Neuropharmacology* **2016**, *101*, 137-145.
- (74) Kolosov, A., Aurini, L., Williams, E.D., Cooke, I., Goodchild, C.S. Intravenous injection of Leconotide, an ω -conotoxin: Synergistic antihyperalgesic effects with morphine in a rat model of bone. *Pain Med.* **2011**, *12*, 923-941.
- (75) Clark, R.J., Jensen, J., Nevin, S.T., Callaghan, B.P., Adams, D.J., Craik, D.J. The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 6545-6548.
- (76) Adams, D.J., Berecki, G. Mechanisms of conotoxin inhibition of N-type (Cav2.2) calcium channels. *Biochim. Biophys. Acta* **2013**, *1828*, 1619-1628.
- (77) Carstens, B.B., Berecki, G., Daniel, J.T., Lee, H.S., Jackson, K.A.V, Tae, H.S., Sadeghi, M., Castro, J., Donnell, T.O., Deiteren, A., Brierley, S.M., David, J., Adams, D.J., Clark, R.J. Structure-activity studies of cysteine-rich α -conotoxins that inhibit high voltage-activated calcium channels via GABA_B receptor activation reveal a minimal functional motif *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 4692-4696.
- (78) Knapp, O., McArthur, J.R., Adams, D.J. Conotoxins targeting neuronal voltage-gated sodium channel subtypes: Potential analgesics? *Toxins* **2012**, *4*, 1236-1260.
- (79) Durek, T., Craik, D.J. Therapeutic conotoxins: A US patent literature survey. *Expert Opin. Ther. Pat.* **2015**, *3776*, 1-15.
- (80) Zhang, M.M., Fiedler, B., Green, B.R., Catlin, P., Watkins, M., Garrett, J.E., Smith, B.J., Yoshikami, D., Olivera, B.M., Bulaj, G. Structural and functional diversities among μ -conotoxins targeting TTX-resistant sodium channels. *Biochemistry* **2006**, *45*, 3723-3732.
- (81) Goldin, A. L. Evolution of voltage-gated Na⁺ channels. *J. Exp. Biol.* **2002**, *205*, 575-584.
- (82) Sands, Z., Grottesi, A., Sansom, M.S.P. Voltage-gated ion channels. *Curr. Biol.* **2005**, *15*, R44-R47.
- (83) Hodgkin, A.L., Huxley, A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **1952**, *117*, 500-544.
- (84) Dutertre, S., Lewis, R.J. Use of venom peptides to probe ion channel structure and function. *J. Biol. Chem.* **2010**, *285*, 13315-13320.
- (85) Catterall, W.A. Structure and function of voltage-gated ion channels. *Annu. Rev. Biochem.* **1995**, *64*, 493-531.
- (86) Catterall, W.A. From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. *Neuron*. **2000**, *26*, 13-25.
- (87) Wood, J.N., Abrahamsen, B., Baker, M.D., Boorman, J.D., Donier, E., Drew, L.J., Nassar, M.A., Okuse, K., Seereeram, A., Stirling, C.L., Zhao, J. Ion channel activities implicated in pathological pain. *Novartis Found Symp.* **2004**, *261*, 32-40.

9. References

- (88) Waxman, S.G., Dib-Hajj, S., Cummins, T.R., Black, J.A. Sodium channels and pain. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 7635-7639.
- (89) Zhang, M.M., Green, B.R., Catlin, P., Fiedler, B., Azam, L., Chadwick, A., Terlau, H., McArthur, J.R., French, R.J., Gulyas, J., Rivier, J.E., Smith, B.J., Norton, R.S., Olivera, B.M., Yoshikami, D., Bulaj, G. Structure/function characterization of micro-conotoxin KIIIA, an analgesic, nearly irreversible blocker of mammalian neuronal sodium channels. *J. Biol. Chem.* **2007**, *282*, 30699-30706.
- (90) Wilson, M.J., Yoshikami, D., Azam, L., Gajewiak, J., Olivera, B.M., Bulaj, G. μ -conotoxins that differentially block sodium channels Nav1.1 through 1.8 identify those responsible for action potentials in sciatic nerve. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10302-10307.
- (91) Yao, S., Zhang, M.M., Yoshikami, D., Azam, L., Olivera, B.M., Bulaj, G., Norton, R.S. Structure, dynamics, and selectivity of the sodium channel blocker μ -conotoxin SIIIA. *Biochemistry* **2008**, *47*, 10940-10949.
- (92) van der Haegen, A., Peigneur, S., Tytgat, J. Importance of position 8 in μ -conotoxin KIIIA for voltage-gated sodium channel selectivity. *FEBS J.* **2011**, *278*, 3408-3418.
- (93) McArthur, J.R., Singh, G., McMaster, D., Winkfein, R., Tieleman, D.P., French, R.J. Interactions of key charged residues contributing to selective block of neuronal sodium channels by μ -conotoxin KIIIA. *Mol. Pharmacol.* **2011**, *80*, 573-584.
- (94) Safo, P., Rosenbaum, T., Shcherbatko, A., Choi, D.Y., Han, E., Toledo-Aral, J.J., Olivera, B.M., Brehm, P., Mandel, G. Distinction among neuronal subtypes of voltage-activated sodium channels by μ -conotoxin PIIAA. *J. Neurosci.* **2000**, *20*, 76-80.
- (95) Schroeder, C.I., Ekberg, J., Nielsen, K.J., Adams, D., Loughnan, M.L., Thomas, L., Adams, D.J., Alewood, P.F., Lewis, R.J. Neuronally micro-conotoxins from *Conus striatus* utilize an α -helical motif to target mammalian sodium channels. *J. Biol. Chem.* **2008**, *283*, 21621-21628.
- (96) Tietze, A.A., Tietze, D., Ohlenschläger, O., Leipold, E., Ullrich, F., Kühl, T., Mischo, A., Buntkowsky, G., Görlach, M., Heinemann, S.H., Imhof, D. Structurally diverse μ -conotoxin PIIAA isomers block sodium channel Na(V)1.4. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 4058-4061.
- (97) Leipold, E., Markgraf, R., Miloslavina, A.A., Kijas, M., Schirmeyer, J., Imhof, D., Heinemann, S.H. Molecular determinants for the subtype specificity of μ -conotoxin SIIIA targeting neuronal voltage-gated sodium channels. *Neuropharmacology* **2011**, *61*, 105-111.
- (98) Leipold, E., Hanson-Kahn, A., Frick, M., Gong, P., Bernstein, J.A., Voigt, M., Katona, I., Goral, R.O., Altmüller, J., Nürnberg, P., Weis, J., Hübner, C.A., Heinemann, S.H., Kurth, I. Cold-aggravated pain in humans caused by a hyperactive Nav1.9 channel mutant. *Nat. Commun.* **2015**, *6*, 10049.
- (99) Kaas, Q., Craik, D.J. Bioinformatics-aided venomomics. *Toxins* **2015**, *7*, 2159-2187.
- (100) Korkosh, V.S., Zhorov, B.S., Tikhonov, D.B. Folding similarity of the outer pore region in prokaryotic and eukaryotic sodium channels revealed by docking of conotoxins GIIIA, PIIAA, and KIIIA in a NavAb-based model of Nav1.4. *J. Gen. Physiol.* **2014**, *144*, 231-244.
- (101) Mahdavi, S., Kuyucak, S. Systematic study of binding of μ -conotoxins to the sodium channel Nav1.4. *Toxins* **2014**, *6*, 3454-3470.
- (102) Patel, D., Mahdavi, S., Kuyucak, S. Computational study of binding of μ -conotoxin GIIIA to bacterial sodium channels NavAb and NavRh. *Biochemistry* **2016**, *55*, 1929-1938.
- (103) Andreu, D., Albericio, F., Solé, N.A., Munson, M.C., Ferrer, M., Barany, G. Formation of disulfide bonds in synthetic peptides and proteins. *Methods Mol. Biol.* **1994**, *35*, 91-169.
- (104) Moroder, L., Besse, D., Musiol, H.J., Rudolph-Böhner, S., Siedler, F. Oxidative folding of cystine-rich peptides vs regioselective cysteine pairing strategies. *Biopolymers* **1996**, *40*, 207-234.
- (105) Moroder, L., Musiol, H.J., Schaschke, N., Chen, L., Hargittai, B., Barany, G. in: *Houben-Weyl, Methods of Organic Chemistry, Synthesis of Peptides and Peptidomimetics*; Goodman, M., Felix, A., Moroder, L., Toniolo, C., (Eds.) Georg Thieme Verlag, Stuttgart, E22; **2001**.
- (106) Postma, T.M., Albericio, F. Disulfide formation strategies in peptide synthesis. *Eur. J. Org. Chem.* **2014**, *17*, 3519-3530.

- (107) Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223-230.
- (108) Anfinsen, C.B., Scheraga, H. Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* **1975**, *29*, 205-300.
- (109) Chang, J.Y. Diverse pathways of oxidative folding of disulfide proteins: Underlying causes and folding models. *Biochemistry* **2011**, *50*, 3414-3431.
- (110) Weissman, J., Kim, P. A kinetic explanation for the rearrangement pathway of BPTI folding. *Nat. Struct. Mol. Biol.* **1995**, *2*, 1123-1130.
- (111) Olga, A. Synthetic cystine-knot miniproteins - Valuable scaffolds for polypeptide engineering. *Adv. Exp. Med. Biol.* **2016**, *917*, 121-144.
- (112) Fersht, A.R. From the first protein structures to our current knowledge of protein folding: Delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 650-654.
- (113) Muttenthaler, M., Nevin, S.T., Grishin, A.A., Ngo, S.T., Choy, P.T., Daly, N.L., Hu, S.H., Armishaw, C.J., Wang, C.I., Lewis, R.J., Martin, J.L., Noakes, P.G., Craik, D.J., Adams, D.J., Alewood, P.F. Solving the α -conotoxin folding problem: Efficient selenium-directed on-resin generation of more potent and stable nicotinic acetylcholine receptor antagonists. *J. Am. Chem. Soc.* **2010**, *132*, 3514-3522.
- (114) Gehrmann, J., Alewood, P.F., Craik, D.J. Structure determination of the three disulfide bond isomers of α -conotoxin GI: A model for the role of disulfide bonds in structural stability. *J. Mol. Biol.* **1998**, *278*, 401-415.
- (115) Dutton, J.L., Bansal, P.S., Hogg, R.C., Adams, D.J., Alewood, P.F., Craik, D.J. A new level of conotoxin diversity, a non-native disulfide bond connectivity in α -conotoxin AulB reduces structural definition but increases biological activity. *J. Biol. Chem.* **2002**, *277*, 48849-48857.
- (116) Khoo, K.K., Gupta, K., Green, B.R., Zhang, M.M., Watkins, M., Olivera, B.M., Balaram, P., Yoshikami, D., Bulaj, G., Norton, R.S. Distinct disulfide isomers of μ -conotoxins KIIIA and KIIIB block voltage-gated sodium channels. *Biochemistry* **2012**, *51*, 9826-9835.
- (117) Metanis, N., Hilvert, D. Harnessing selenocysteine reactivity for oxidative protein folding. *Chem. Sci.* **2015**, *6*, 322-325.
- (118) Craik, D.J. Protein folding: Turbo-charged crosslinking. *Nat. Chem.* **2012**, *4*, 600-602.
- (119) Walewska, A., Zhang, M.M., Skalicky, J.J., Yoshikami, D., Olivera, B.M., Bulaj, G. Integrated oxidative folding of cysteine/selenocysteine containing peptides: Improving chemical synthesis of conotoxins. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 2221-2224.
- (120) Zhou, M., Wang, L., Wu, Y., Liu, J., Sun, D., Zhu, X., Feng, Y., Qin, M., Chen, S., Xu, A. Soluble expression and sodium channel activity of It16a, a novel framework XVI conotoxin from the M-Superfamily. *Toxicon* **2015**, *98*, 5-11.
- (121) Zhu, X., Bi, J., Yu, J., Li, X., Zhang, Y., Zhangsun, D., Luo, S. Recombinant expression and characterization of α -conotoxin LvIA in *Escherichia coli*. *Mar. Drugs* **2016**, *14*, 1-15.
- (122) Akaji, K., Fujino, K., Tatsumi, T., Kiso, Y. Regioselective double disulfide formation using silylchloride-sulfoxide system. *Tetrahedron Lett.* **1992**, *33*, 1073-1076.
- (123) Szabo, I., Schlosser, G., Hudecz, F., Mezo, G. Disulfide bond rearrangement during regioselective oxidation in PhS(O)Ph/CH₃SiCl₃ mixture for the synthesis of α -conotoxin GI. *Biopolymers* **2007**, *88*, 20-28.
- (124) Mochizuki, M., Tsuda, S., Tanimura, K., Nishiuchi, Y. Regioselective formation of multiple disulfide bonds with the aid of postsynthetic S-tritylation. *Org. Lett.* **2015**, *17*, 2202-2205.
- (125) Peigneur, S., Paolini-Bertrand, M., Gärtner, H., Biass, D., Violette, A., Stöcklin, R., Favreau, P., Tytgat, J., Hartley, O. δ -conotoxins synthesized using an acid-cleavable solubility tag approach reveal key structural determinants for Nav subtype selectivity. *J. Biol. Chem.* **2014**, *289*, 35341-35350.
- (126) Cuthbertson, A., Indrevoll, B. Regioselective formation, using orthogonal cysteine protection, of an α -conotoxin dimer peptide containing four disulfide bonds. *Org. Lett.* **2003**, *5*, 2955-2957.
- (127) Dekan, Z., Mobli, M., Pennington, M.W., Fung, E., Nemeth, E., Alewood, P.F. Total synthesis of human

9. References

- hepcidin through regioselective disulfide-bond formation by using the safety-catch cysteine protecting group 4,4'-Dimethylsulfinylbenzhydryl. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 2931-2934.
- (128) Hargittai, B., Barany, G. Controlled syntheses of natural and disulfide-mispaired regioisomers of α -conotoxin SI. *J. Pept. Res.* **1999**, *54*, 468-479.
- (129) Creighton, T.E. The disulfide folding pathway of BPTI. *Science* **1992**, *256*, 111-114.
- (130) Creighton, T.E. Experimental studies of protein folding and unfolding. *Prog. Biophys. Mol. Biol.* **1978**, *33*, 231-297.
- (131) Bartlett, A.I., Radford, S.E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **2009**, *16*, 582-588.
- (132) Schultz, C.P. Illuminating folding intermediates. *Nat. Struct. Biol.* **2000**, *7*, 7-10.
- (133) Dobson, C.M. Protein folding and misfolding. *Nature* **2003**, *426*, 884-890.
- (134) Carrell, R.W., Lomas, D.A. Conformational disease. *Lancet* **1997**, *350*, 134-138.
- (135) Dill, K.A., MacCallum, J.L. The protein-folding problem, 50 years on. *Science* **2012**, *338*, 1042-1046.
- (136) Dill, K.A., Ozkan, S.B., Weikl, T.R., Chodera, J.D., Voelz, V.A. The protein folding problem: When will it be solved? *Curr. Opin. Struct. Biol.* **2007**, *17*, 342-346.
- (137) Muñoz, V. in: *Protein folding, misfolding and aggregation: Classical themes and novel approaches*. The Royal Society of Chemistry, Cambridge, **2008**.
- (138) Liu, S., Ji, X., Tao, Y., Tan, D., Zhang, K.Q., Fu, Y.X. in: *Protein Engineering; Protein folding, binding and energy landscape: A synthesis*. (Eds. Kaumaya P.) InTech, Rijeka **2012**.
- (139) Chatrenet, B., Chang, J.Y. The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J. Biol. Chem.* **1993**, *268*, 20988-20996.
- (140) Cemazar, M., Zaharieva, S., Lopez, J.J., Carugo, O., Jones, J.A., Hore, P.J., Pongor, S. Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5754-5759.
- (141) Chang, J.Y., Lu, B.Y., Lin, C.C.J., Yu, C. Fully oxidized scrambled isomers are essential and predominant folding intermediates of Cardiotoxin-III. *FEBS Lett.* **2006**, *580*, 656-660.
- (142) Böhm, M., Bäuml, C.A., Hardes, K., Steinmetzer, T., Roeser, D., Schaub, Y., Than, M.E., Biswas, A., Imhof, D. Novel insights into structure and function of Factor XIIIa-inhibitor Tridegin. *J. Med. Chem.* **2014**, *57*, 10355-10365.
- (143) Bulaj, G., Olivera, B.M. Folding of conotoxins: Formation of the native disulfide bridges during chemical synthesis and biosynthesis of *Conus* peptides. *Antioxid. Redox Signal.* **2008**, *10*, 141-155.
- (144) Price-Carter, M., Gray, W.R., Goldenberg, D.P. Folding of ω -conotoxins. 1. Efficient disulfide-coupled folding of mature sequences *in vitro*. *Biochemistry* **1996**, *35*, 15537-15546.
- (145) Steiner, A.M., Bulaj, G. Optimization of oxidative folding methods for cysteine-rich peptides: A study of conotoxins containing three disulfide bridges. *J. Pept. Sci.* **2011**, *17*, 1-7.
- (146) Domínguez de María, P. in: *Ionic Liquids in Biotransformations and Organocatalysis: Solvents and Beyond*. John Wiley & Sons, Weinheim, **2012**.
- (147) Ghandi, K. A review of Ionic Liquids, their limits and applications. *Green Sustain. Chem.* **2014**, *4*, 44-53.
- (148) Thomas, P.A., Marvey, B.B. Room temperature Ionic Liquids as green solvent alternatives in the metathesis of oleochemical feedstocks. *Molecules* **2016**, *21*, 1-16.
- (149) Kuchenbuch, A., Giernoth, R. Ionic Liquids beyond simple solvents: Glimpses at the state of the art in organic chemistry. *Chemistry Open* **2015**, *4*, 677-681.
- (150) Welton, T. Solvents and sustainable chemistry. *Proc. Math. Phys. Eng. Sci.* **2015**, *471*, 1-26.
- (151) van Rantwijk, F., Sheldon, R.A. Biocatalysis in Ionic Liquids. *Chem. Rev.* **2007**, *107*, 2757-2785.

- (152) Sheldon, R.A. Biocatalysis and biomass conversion in alternative reaction media. *Chemistry* **2016**, *22*, 1-17.
- (153) Stark, A. Ionic Liquids in the biorefinery: A critical assessment of their potential. *Energy Environ. Sci.* **2011**, *4*, 19-32.
- (154) Çinar, S., Schulz, M.D., Oyola-Reynoso, S., Bwambok, D.K., Gathiaka, S.M., Thuo, M. Application of Ionic Liquids in pot-in-pot reactions. *Molecules* **2016**, *21*, 1-14.
- (155) Irvine, J.T.S. A perspective on liquid salts for energy and materials. *Faraday Discuss.* **2016**, *190*, 551-559.
- (156) Plaquevent, J.C., Levillain, J., Guillen, F., Malhiac, C., Gaumont, A.C. Ionic Liquids: New targets and media for α -amino acid and peptide chemistry. *Chem. Rev.* **2008**, *108*, 5035-5060.
- (157) Zhao, H. Protein stabilization and enzyme activation in Ionic Liquids: Specific ion effects. *J. Chem. Technol. Biotechnol.* **2016**, *91*, 25-50.
- (158) Hofmeister, F. Zur Lehre von der Wirkung der Salze. Zweite Mitteilung. *Arch. Exp. Pathol. Pharmacol.* **1888**, *24*, 247-260.
- (159) Constantinescu, D., Weingärtner, H., Herrmann, C. Protein denaturation by Ionic Liquids and the Hofmeister series: A case study of aqueous solutions of Ribonuclease A. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 8887-8889.
- (160) Zhang, Y., Cremer, P.S. Interactions between macromolecules and ions: The Hofmeister series. *Curr. Opin. Chem. Biol.* **2006**, *10*, 658-663.
- (161) Zhao, H., Olubajo, O., Song, Z., Sims, A.L., Person, T.E., Lawal, R.A., Holley, L.A. Effect of kosmotropicity of Ionic Liquids on the enzyme stability in aqueous solutions. *Bioorg. Chem.* **2006**, *34*, 15-25.
- (162) Zhao, H., Campbell, S.M., Jackson, L., Song, Z., Olubajo, O. Hofmeister series of Ionic Liquids: Kosmotropic effect of Ionic Liquids on the enzymatic hydrolysis of enantiomeric phenylalanine methyl ester. *Tetrahedron: Asymmetry* **2006**, *17*, 377-383.
- (163) Vallette, H., Ferron, L., Coquerel, G., Gaumont, A.C., Plaquevent, J.C. Peptide synthesis in room temperature Ionic Liquids. *Tetrahedron Lett.* **2004**, *45*, 1617-1619.
- (164) Chen, L., Zheng, M.F., Zhou, Y., Liu, H., Jiang, H.L. Ionic-Liquid-supported total synthesis of Sansalvamide A peptide. *Synth. Commun.* **2008**, *38*, 239-248.
- (165) He, X., Chan, T.H. Structurally-defined imidazolium-type ionic oligomers as soluble/solid support for peptide synthesis. *Org. Lett.* **2007**, *9*, 2681-2684.
- (166) Miloslavina, A.A., Leipold, E., Kijas, M., Stark, A., Heinemann, S.H., Imhof, D. A room temperature Ionic Liquid as convenient solvent for the oxidative folding of conopeptides. *J. Pept. Sci.* **2009**, *15*, 72-77.
- (167) Koel, M. Ionic Liquids in chemical analysis. *Crit. Rev. Anal. Chem.* **2005**, *35*, 177-192.
- (168) Mrestani-Klaus, C., Richardt, A., Wespe, C., Stark, A., Humpfer, E., Bordusa, F. Structural studies on Ionic Liquid/water/peptide systems by HR-MAS NMR spectroscopy. *Chem. Phys. Chem.* **2012**, *13*, 1836-1844.
- (169) Tietze, A.A., Bordusa, F., Giernoth, R., Imhof, D., Lenzer, T., Maaß, A., Mrestani-Klaus, C., Neundorff, I., Oum, K., Reith, D., Stark, A. On the nature of interactions between Ionic Liquids and small amino-acid-based biomolecules. *Chem. Phys. Chem.* **2013**, *14*, 4044-4064.
- (170) Khemchyan, L.L., Khokhlova, E.A., Seitkalieva, M.M., Ananikov, V.P. Efficient sustainable tool for monitoring chemical reactions and structure determination in Ionic Liquids by ESI-MS. *Chemistry Open* **2013**, *2*, 208-214.
- (171) Guo, M., Zhai, Y., Guo, C., Liu, Y., Tang, D., Pan, Y. A new strategy to determine the protein mutation site using Matrix-assisted laser desorption ionization in-source decay: Derivatization by Ionic Liquid. *Anal. Chim. Acta* **2015**, *865*, 31-38.
- (172) Debois, D., Smargiasso, N., Demeure, K., Asakawa, D., Zimmermann, T.A., Quinton, L., Pauw, E.D. MALDI in-source decay, from sequencing to imaging. *Top. Curr. Chem.* **2012**, *11*, 13-35.
- (173) The Royal Swedish Academy of Sciences. Mass Spectrometry (MS) and Nuclear Magnetic Resonance

9. References

- (NMR) applied to biological macromolecules. *Adv. Inf. Nobel Prize Chem.* **2002**, 1–13.
- (174) Drenth, J. in: Principles of Protein X-Ray Crystallography. Springer Verlag, Berlin, **2007**.
- (175) Wüthrich, K. Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **1990**, *265*, 22059-22062.
- (176) Poppe, L., Hui, J.O., Ligutti, J., Murray, J.K., Schnier, P.D. PADLOC: A powerful tool to assign disulfide bond connectivities in peptides and proteins by NMR spectroscopy. *Anal. Chem.* **2012**, *84*, 262-266.
- (177) Anand, P., Grigoryan, A., Bhuiyan, M.H., Ueberheide, B., Russell, V., Quinonez, J., Moy, P., Chait, B.T., Poget, S.F., Holford, M. Sample limited characterization of a novel disulfide-rich venom peptide toxin from Terebrid marine snail *Terebra variegata*. *PLoS One* **2014**, *9*, e94122.
- (178) Aebersold, R., Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198-207.
- (179) Edman, P., Begg, G. A protein sequenator. *Eur. J. Biochem.* **1967**, *1*, 80-91.
- (180) Shu, Q., Huang, R., Liang, S. Assignment of the disulfide bonds of Huwentoxin-II by Edman degradation sequencing and stepwise thiol modification. *Eur. J. Biochem.* **2001**, *268*, 2301-2307.
- (181) Haniu, M., Acklin, C., Kenney, W.C., Rohde, M.F. Direct assignment of disulfide bonds by Edman degradation of selected peptide fragments. *Int. J. Pept. Protein Res.* **1994**, *43*, 81-86.
- (182) Verdes, A., Anand, P., Gorson, J., Jannetti, S., Kelly, P., Leffler, A., Simpson, D., Ramrattan, G., Holford, M. From mollusks to medicine: A venomomics approach for the discovery and characterization of therapeutics from Terebridae peptide toxins. *Toxins* **2016**, *8*, 1-30.
- (183) Bodnar, W.M., Blackburn, R.K., Krise, J.M., Moseley, M.A. Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 971-979.
- (184) Jenkins, R., Duggan, J.X., Aubry, A.F., Zeng, J., Lee, J.W., Cojocar, L., Dufield, D., Garofolo, F., Kaur, S., Schultz, G., Xu, K., Yang, Z., Yu, J., Zhang, Y.J., Vazvaei, F. Recommendations for validation of LC-MS/MS bioanalytical methods for protein biotherapeutics. *AAPS J.* **2015**, *17*, 1-16.
- (185) Kuklenyik, Z., Boyer, A.E., Lins, R., Quinn, C.P., Gallegos-Candela, M., Woolfitt, A., Pirkle, J.L., Barr, J.R. Comparison of MALDI-TOF-MS and HPLC-ESI-MS/MS for endopeptidase activity-based quantification of Anthrax lethal factor in serum. *Anal. Chem.* **2011**, *83*, 1760-1765.
- (186) Nesvizhskii, A.I., Vitek, O., Aebersold, R. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* **2007**, *4*, 787-797.
- (187) Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A.V., Meehan, M.J., Liu, W.T., Crüsemann, M., Boudreau, P.D., Esquenazi, E., Sandoval-Calderón, M., Kersten, R.D., Pace, L.A., Quinn, R.A., Duncan, K.R., Hsu, C.C., Floros, D.J., Gavilan, R.G., Kleigrew, K., Northen, T., Dutton, R.J., Parrot, D., Carlson, E.E., Aigle, B., Michelsen, C.F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy, B.T., Gerwick, L., Liaw, C.C., Yang, Y.L., Humpf, H.U., Maansson, M., Keyzers, R.A., Sims, A.C., Johnson, A.R., Sidebottom, A.M., Sedito, B.E., Klitgaard, A., Larson, C.B., Boya, P.C.A., Torres-Mendoza, D., Gonzalez, D.J., Silva, D.B., Marques, L.M., Demarque, D.P., Pociute, E., O'Neill, E.C., Briand, E., Helfrich, E.J.N., Granatosky, E.A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J.J., Zeng, Y., Vorholt, J.A., Kurita, K.L., Charusanti, P., McPhail, K.L., Nielsen, K.F., Vuong, L., Elfeki, M., Traxler, M.F., Engene, N., Koyama, N., Vining, O.B., Baric, R., Silva, R.R., Mascuch, S.J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P.G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A.M.C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B.M., Almaliti, J., Allard, P.M., Phapale, P., Nothias, L.F., Alexandrov, T., Litaudon, M., Wolfender, J.L., Kyle, J.E., Metz, T.O., Peryea, T., Nguyen, D.T., VanLeer, D., Shinn, P., Jadhav, A., Müller, R., Waters, K.M., Shi, W., Liu, X., Zhang, L., Knight, R., Jensen, P.R., Palsson, B.O., Poglian, K., Linington, R.G., Gutiérrez, M., Lopes, N.P., Gerwick, W.H., Moore, B.S., Dorrestein, P.C., Bandeira, N. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* **2016**, *34*, 828-837.
- (188) Murad, W., Singh, R., Yen, T.Y. An efficient algorithmic approach for mass spectrometry-based disulfide connectivity determination using multi-ion analysis. *BMC Bioinformatics* **2011**, *12*, S1-12.
- (189) Gorman, J.J., Wallis, T.P., Pitt, J.J. Protein disulfide bond determination by mass spectrometry. *Mass Spectrom. Rev.* **2002**, *21*, 183-216.

- (190) Goyder, M.S., Rebeaud, F., Pfeifer, M.E., Kálmán, F. Strategies in mass spectrometry for the assignment of Cys-Cys disulfide connectivities in proteins. *Expert Rev. Proteomics* **2013**, *10*, 489-501.
- (191) Gray, W.R. Disulfide structures of highly bridged peptides : A new strategy for analysis. *Protein Sci.* **1993**, *2*, 1732-1748.
- (192) Yen, T.Y., Yan, H., Macher, B.A. Characterizing closely spaced, complex disulfide bond patterns in peptides and proteins by liquid chromatography/electrospray ionization tandem mass spectrometry. *J. Mass Spectrom.* **2002**, *37*, 15-30.
- (193) Borges, C.R., Sherma, N.D. Techniques for the analysis of cysteine sulfhydryls and oxidative protein folding. *Antioxid. Redox Signal.* **2014**, *21*, 511-531.
- (194) Fenn, J.B. Electrospray wings for molecular elephants (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3871-3894.
- (195) Banerjee, S., Mazumdar, S. Electrospray ionization mass spectrometry: A technique to access the information beyond the molecular weight of the analyte. *Int. J. Anal. Chem.* **2012**, *2012*, 1-40.
- (196) Dutertre, S., Undheim, E.A.B., Pineda, S.S., Jin, A.H., Laverigne, V., Fry, B.G., Lewis, R.J., Alewood, P.F., King, G.F. in: Venoms to drugs: Venom as a source for the development of human therapeutics. The Royal Society of Chemistry, Cambridge **2015**, 80-96.
- (197) Aebersold, R., Goodlett, D.R. Mass spectrometry in proteomics. *Chem. Rev.* **2001**, *101*, 269-295.
- (198) Lingdong, Q., Miao, L. CID, ETD and HCD fragmentation to study protein post-translational modifications. *Mod. Chem. Appl.* **2013**, *1*, 1-2.
- (199) Echterbille, J., Gilles, N., Pauw, E.D. Ion mobility mass spectrometry as a potential tool to assign disulfide bonds arrangements in peptides with multiple disulfide bridges. *Anal. Chem.* **2013**, *85*, 4405-4413.
- (200) Tanaka, K. The origin of macromolecule ionization by laser irradiation (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3860-3870.
- (201) Schnaible, V., Wefing, S., Resemann, A., Suckau, D., Bücken, A., Wolf-Kümmeth, S., Hoffmann, D. Screening for disulfide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS. *Anal. Chem.* **2002**, *74*, 4980-4988.
- (202) Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J., Holle, A.A Novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* **2003**, *376*, 952-965.
- (203) Quinton, L., Demeure, K., Dobson, R., Gilles, N., Gabelica, V., Pauw, E.D. New method for characterizing highly disulfide-bridged peptides in complex mixtures: Application to toxin identification from crude venoms. *J. Proteome Res.* **2007**, *6*, 3216-3223.
- (204) Zhao, L., Almaraz, R.T., Xiang, F., Hedrick, J.L., Franz, A.H. Gas-phase scrambling of disulfide bonds during matrix-assisted laser desorption/ionization mass spectrometry analysis. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1603-1616.
- (205) Nair, S.S., Nilsson, C.L., Emmett, M.R., Schaub, T.M., Gowd, K.H., Thakur, S.S., Krishnan, K.S., Balaram, P., Marshall, A.G. *De novo* sequencing and disulfide mapping of a bromotryptophan-containing conotoxin by Fourier Transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2006**, *78*, 8082-8088.
- (206) Crecelius, A.C., Schubert, U.S., von Eggeling, F. MALDI mass spectrometric imaging meets "omics": Recent advances in the fruitful marriage. *Analyst.* **2015**, *140*, 5806-5820.
- (207) Gupta, K., Kumar, M., Balaram, P. Disulfide bond assignments by mass spectrometry of native natural peptides: Cysteine pairing in disulfide bonded conotoxins. *Anal. Chem.* **2010**, *82*, 8313-8319.
- (208) Cruz, L.J., Gray, W.R., Olivera, B.M., Zeikus, R.D., Kerr, L., Yoshikami, D., Moczydlowski, E. *Conus geographus* toxins that discriminate between neuronal and muscle sodium channels. *J. Biol. Chem.* **1985**, *260*, 9280-9288.
- (209) Shon, K.J., Olivera, B.M., Watkins, M., Jacobsen, R.B., Gray, W.R., Floresca, C.Z., Cruz, L.J., Hillyard, D.R., Brink, A., Terlau, H., Yoshikami, D. μ -conotoxin PIIIA, a new peptide for discriminating among Tetrodotoxin-sensitive Na channel subtypes. *J. Neurosci.* **1998**, *18*, 4473-4481.
- (210) Volpon, L., Lamthanh, H., Barbier, J., Gilles, N., Molgo, J., Menez, A., Lancelin, J.M. NMR solution

9. References

- structures of δ -conotoxin EVIA from *Conus ermineus* that selectively acts on vertebrate neuronal Na⁺ channels. *J. Biol. Chem.* **2004**, *279*, 21356-21366.
- (211) Shon, K.J., Hasson, A., Spira, M.E., Cruz, L.J., Gray, W.R., Olivera, B.M. δ -conotoxin GmVIA, a novel peptide from the venom of *Conus gloriamaris*. *Biochemistry* **1994**, *33*, 11420-11425.
- (212) Akcan, M., Cao, Y., Chongxu, F., Craik, D.J. The three-dimensional solution structure of mini-M conotoxin BtIIIA reveals a disconnection between disulfide connectivity and peptide fold. *Bioorganic Med. Chem.* **2013**, *21*, 3590-3596.
- (213) Kancherla, A.K., Meesala, S., Jorwal, P., Palanisamy, R., Sikdar, S.K., Sarma, S.P. A disulfide stabilized β -sandwich defines the structure of a new cysteine framework M-superfamily conotoxin. *ACS Chem. Biol.* **2015**, *10*, 1847-1860.
- (214) Saunders, M., Wishnia, A., Kirkwood, J.G. The nuclear magnetic resonance spectrum of Ribonuclease. *J. Am. Chem. Soc.* **1957**, *79*, 3289-3290.
- (215) Wüthrich, K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *J. Biomol. NMR* **2003**, *27*, 13-39.
- (216) Pellecchia, M., Sem, D.S., Wüthrich, K. NMR in drug discovery. *Nat. Rev. Drug Discov.* **2002**, *1*, 211-219.
- (217) Keeler, J. in: Understanding NMR spectroscopy. John Wiley & Sons, Weinheim, **2010**.
- (218) Wüthrich, K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 3340-3363.
- (219) Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687-696.
- (220) Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., Vriend, G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins* **2004**, *57*, 678-683.
- (221) Gutmanas, A., Adams, P.D., Bardiaux, B., Berman, H.M., Case, D.A., Fogh, R.H., Güntert, P., Hendrickx, P.M.S., Herrmann, T., Kleywegt, G.J., Kobayashi, N., Lange, O.F., Markley, J.L., Montelione, G.T., Nilges, M., Ragan, T.J., Schwieters, C.D., Tejero, R., Ulrich, E.L., Velankar, S., Vranken, W.F., Wedell, J.R., Westbrook, J., Wishart, D.S., Vuister, G.W. NMR exchange format: A unified and open standard for representation of NMR restraint data. *Nat. Struct. Mol. Biol.* **2015**, *22*, 433-434.
- (222) O'Donoghue, S.I., Goodsell, D.S., Frangakis, A.S., Jossinet, F., Laskowski, R.A., Nilges, M., Saibil, H.R., Schafferhans, A., Wade, R.C., Westhof, E., Olson, A.J. Visualization of macromolecular structures. *Nat. Methods* **2010**, *7*, S42-S55.
- (223) Kozlikova, B., Krone, M., Lindow, N., Falk, M., Baaden, M., Baum, D., Viola, I., Parulek, J., Hege, H.C. Visualization of biomolecular structures: State of the art. *Eurographics Conf. Vis.* **2015**, 1-21.
- (224) Sussman, J.L., Lin, D.W., Jiang, J.S., Manning, N.O., Prilusky, J., Ritter, O., Abola, E.E. Protein Data Bank (PDB): Database of three-dimensional structural information of biological macromolecules. *Acta Cryst.* **1998**, *54*, 1078-1084.
- (225) Hooft, R.W., Sander, C., Scharf, M., Vriend, G. The PDBFINDER Database: A summary of PDB, DSSP and HSSP information with added value. *Comput. Appl. Biosci.* **1996**, *12*, 525-529.
- (226) Han, Y.H., Wang, Q., Jiang, H., Liu, L., Xiao, C., Yuan, D.D., Shao, X.X., Dai, Q.Y., Cheng, J.S., Chi, C.W. Characterization of novel M-superfamily conotoxins with new disulfide linkage. *FEBS J.* **2006**, *273*, 4972-4982.
- (227) McDougal, O.M., Poulter, C.D. Three-dimensional structure of the mini-M conotoxin mr3a. *Biochemistry* **2004**, *43*, 425-429.
- (228) Mobli, M., King, G.F. NMR methods for determining disulfide-bond connectivities. *Toxicon* **2010**, *56*, 849-854.
- (229) He, Q.Y., He, Q.Z., Deng, X.C., Yao, L., Meng, E., Liu, Z.H., Liang, S.P. ATDB: A uni-database platform for animal toxins. *Nucleic Acids Res.* **2008**, *36*, 293-297.
- (230) Liu, H., Boudreau, M.A., Zheng, J., Whittall, R.M., Austin, P., Roskelley, C.D., Roberge, M., Andersen,

- R.J., Vederas, J.C. Chemical synthesis and biological activity of the neopetrosiamides and their analogues: Revision of disulfide bond connectivity. *J. Am. Chem. Soc.* **2010**, *132*, 1486-1487.
- (231) Boisbouvier, J., Blackledge, M., Sollier, A., Marion, D. Simultaneous setermination of disulphide bridge topology and three-dimensional structure using ambiguous intersulphur distance restraints: Possibilities and limitations. *J. Biomol. NMR* **2000**, *16*, 197-208.
- (232) Walewska, A., Skalicky, J.J., Davis, D.R., Zhang, M.M., Lopez-Vera, E., Watkins, M., Han, T.S., Yoshikami, D., Olivera, B.M., Bulaj, G. NMR-based mapping of disulfide bridges in cysteine-rich peptides: Application to the μ -conotoxin SxIIIA. *J. Am. Chem. Soc.* **2008**, *130*, 14280-14286.
- (233) Wasserscheid, P., Welton, T. in: *Ionic Liquids in Synthesis*. Wiley-VCH, Weinheim, **2007**.
- (234) Tietze, A.A., Heimer, P., Stark, A., Imhof, D. Ionic Liquid applications in peptide chemistry: Synthesis, purification and analytical characterization processes. *Molecules* **2012**, *17*, 4158-4185.
- (235) Böhm, M., Kühl, T., Harges, K., Coch, R., Arkona, C., Schlott, B., Steinmetzer, T., Imhof, D. Synthesis and functional characterization of Tridegin and its analogues: Inhibitors and substrates of Factor XIIIa. *Chem. Med. Chem.* **2012**, *7*, 326-333.
- (236) Hackeng, T.M., Griffin, J.H., Dawson, P.E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068-10073.
- (237) Kühl, T., Chen, M., Teichmann, K., Stark, A., Imhof, D. Ionic Liquid 1-ethyl-3-methylimidazolium acetate: An attractive solvent for native chemical ligation of peptides. *Tetrahedron Lett.* **2014**, *55*, 3658-3662.
- (238) Weingärtner, H. Understanding Ionic Liquids at the molecular level: Facts, problems, and controversies. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 654-670.
- (239) Lesch, V., Heuer, A., Tatsis, V.A., Holm, C., Smiatek, J. Peptides in the presence of queous Ionic Liquids: Tunable co-solutes as denaturants or protectants? *Phys. Chem. Chem. Phys.* **2015**, *17*, 26049-26053.
- (240) Barbier, J., Lamthanh, H., le Gall, F., Favreau, P., Benoit, E., Chen, H., Gilles, N., Ilan, N., Heinemann, S.H., Gordon, D., Menez, A., Molgo, J. A δ -conotoxin from *Conus ermineus* venom inhibits inactivation in vertebrate neuronal Na⁺ channels but not in skeletal and cardiac muscles. *J. Biol. Chem.* **2004**, *279*, 4680-4685.
- (241) Khoo, K.K., Feng, Z.P., Smith, B.J., Zhang, M.M., Yoshikami, D., Olivera, B.M., Bulaj, G., Norton, R.S. Structure of the analgesic μ -conotoxin KIIIA and effects on structure and function of disulfide deletion. *Biochemistry* **2009**, *48*, 1210-1219.
- (242) Han, T.S., Zhang, M.M., Walewska, A., Gruszczynski, P., Robertson, C.R., Cheatham, T.E., Yoshikami, D., Olivera, B.M., Bulaj, G. Structurally minimized μ -conotoxin analogues as sodium channel blockers: Implications for designing conopeptide-based therapeutics. *Chem. Med. Chem.* **2009**, *4*, 406-414.
- (243) Han, P., Wang, K., Dai, X., Cao, Y., Liu, S., Jiang, H., Fan, C., Wu, W., Chen, J. The role of individual disulfide bonds of μ -conotoxin GIIIA in the inhibition of Nav1.4. *Mar. Drugs* **2016**, *14*, 1-9.
- (244) Karampudi, N.B.R., Bahadur, R.P. Layers: A molecular surface peeling algorithm and its applications to analyze protein structures. *Sci. Rep.* **2015**, *5*, 16141.
- (245) Jones, J.H. Abbreviations and symbols in peptide science: A revised guide and commentary. *J. Pept. Sci.* **2006**, *12*, 1-12.

Appendix

Supporting Information

Computational study on μ -conotoxin conformations using molecular dynamics and computer visualization methods

Ajay Abisheck Paul George,^[a] Pascal Heimer,^[a] Raphael Menges,^[b] Veda Thota,^[a] Astrid Maaß,^[c] Jan Hamaekers,^[c] Martin Hofmann-Apitius,^[d,e] Nils Lichtenberg,^[b] Kai Lawonn,^[b] Diana Imhof^{*[a]}

Table of contents

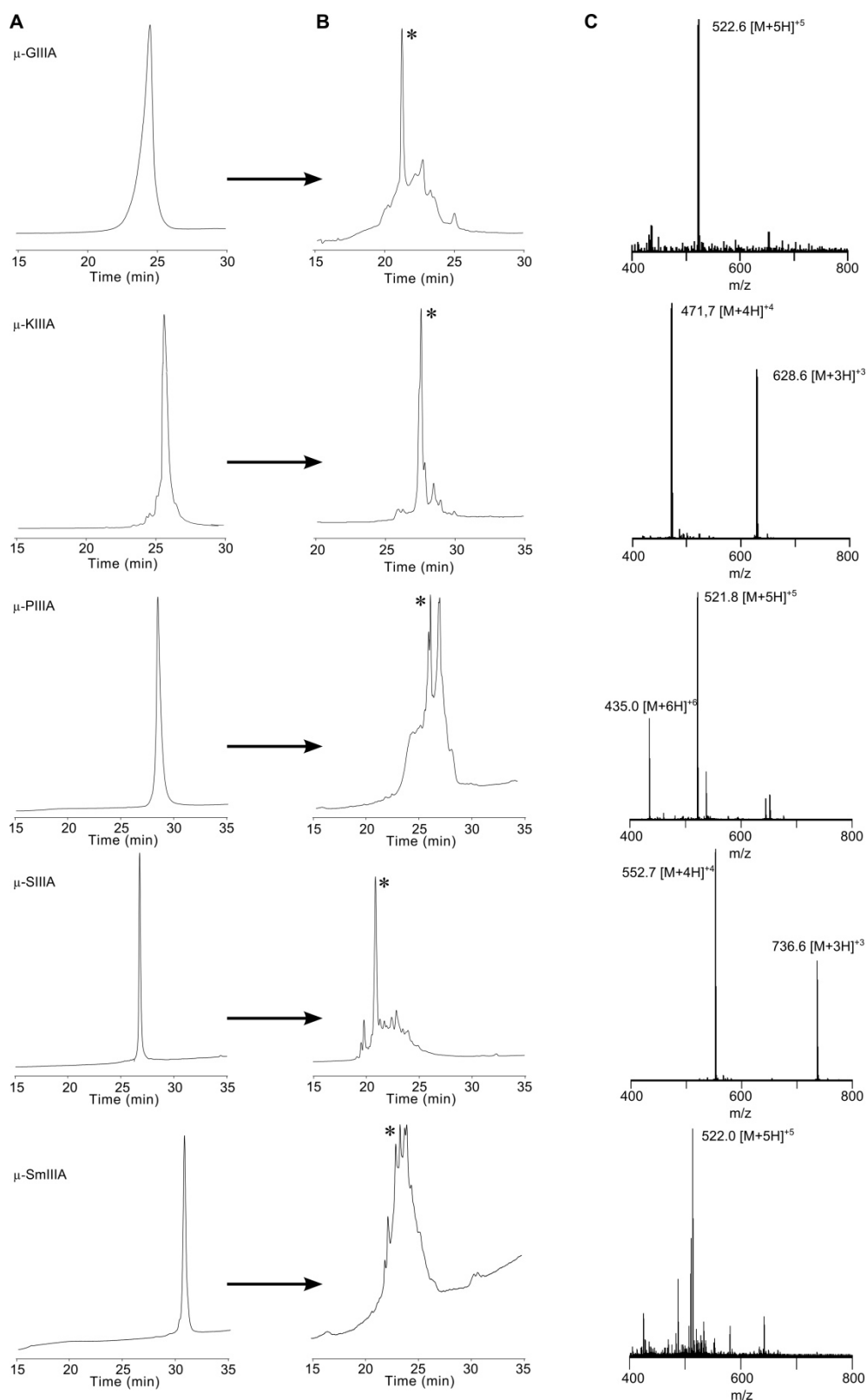
Supporting Figures

S1. HPLC elution profile	2
S2. RMSD, RMSF, and Rg of the five μ -conotoxins	3
S3. MD simulation of μ -PIIIA for 1 μ s	4
S4. Backbone stability of the five μ -conotoxins with 3, 2, and 1 disulfide-bridge	5
S5. Backbone conformation of μ -GIIIA and μ -SmIIIA, helix formation	6
S6. Backbone conformation of μ -PIIIA, loss of helix by disulfide-bond removal	7
S7. VDW Surface representation and SAS of the five μ -conotoxins.	8
S8. Heatmaps of μ -conotoxins KIIIA, PIIIA, and SIIIA over 100 ns simulation	9
S9. Bar chart showing Δ LR for the residues of all five μ -conotoxins	10

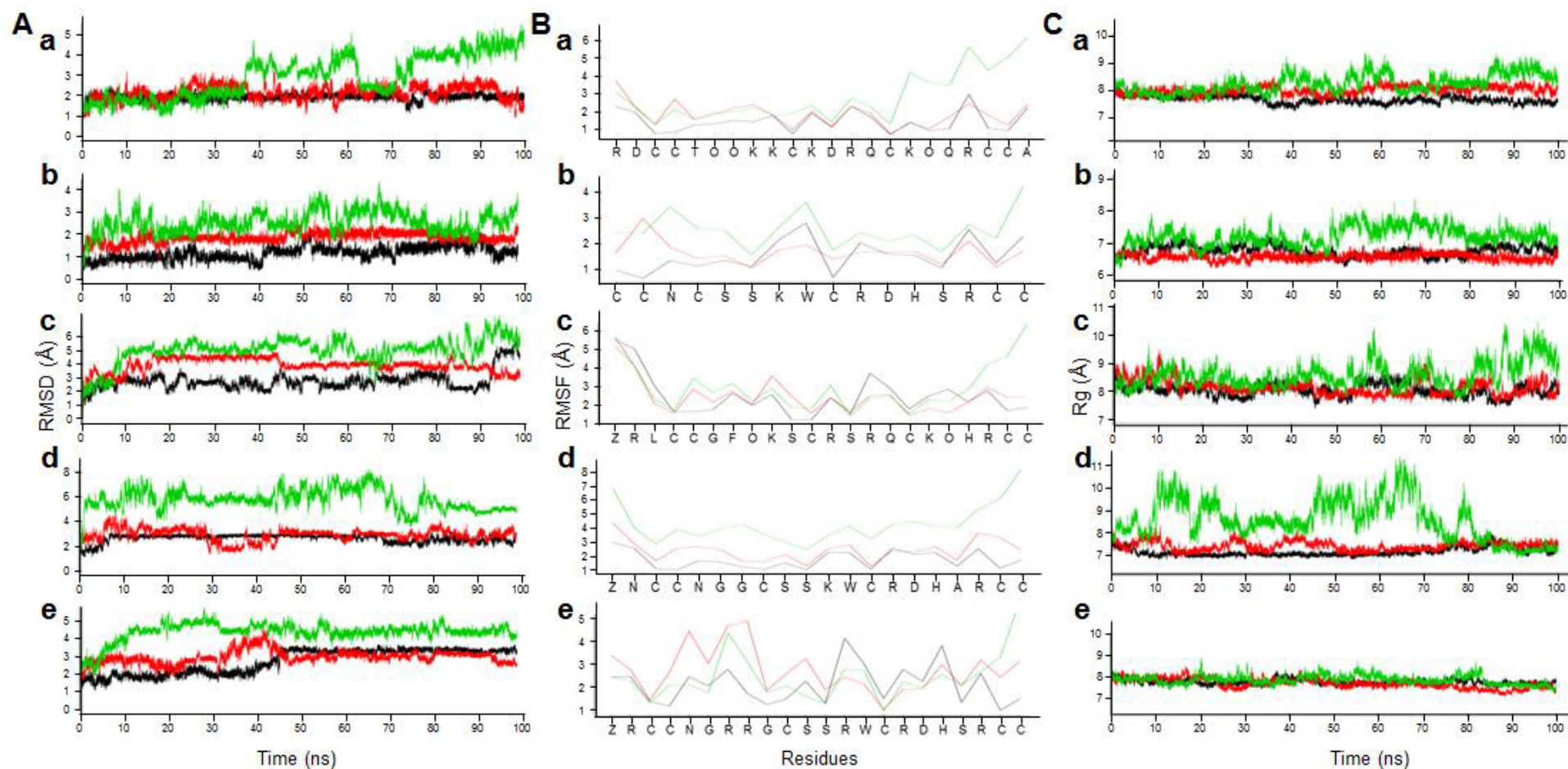
Supporting Tables

S1. Comparison of sequence characteristics of the five μ -conotoxins	11
S2. Comparison of disulfide loop length of the five μ -conotoxins	12
S3. Assessment of stability using RMSD and Rg for 100 ns MD simulation	13
S4. Impact of the C2-C5 and C3-C6 disulfide bridge on the backbone RMSF and the functionally stable residues	14

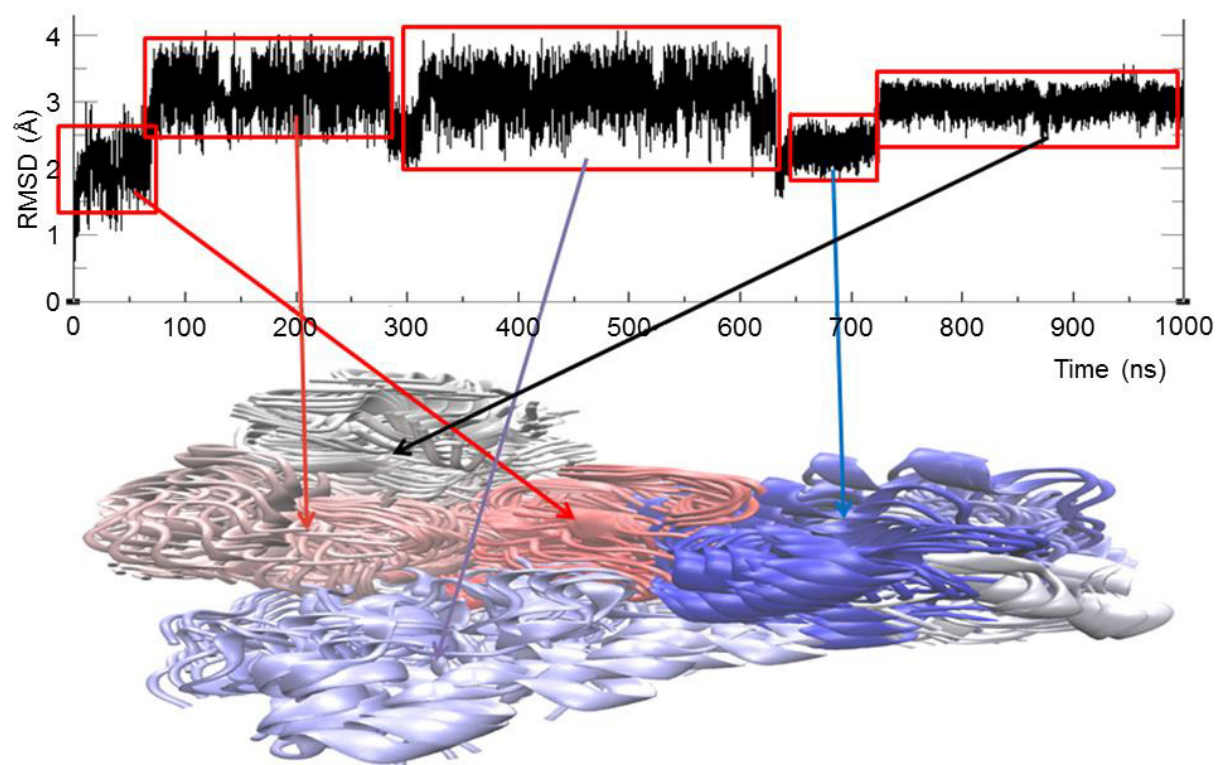
Supporting Figure 1. HPLC elution profiles of μ -conotoxins. Linear, reduced precursors of μ -conotoxins (A), and folded crude mixtures after 1 h (B) with the main product marked with an asterisk. Respective ESI mass spectra of the oxidized peptides are shown in (C).



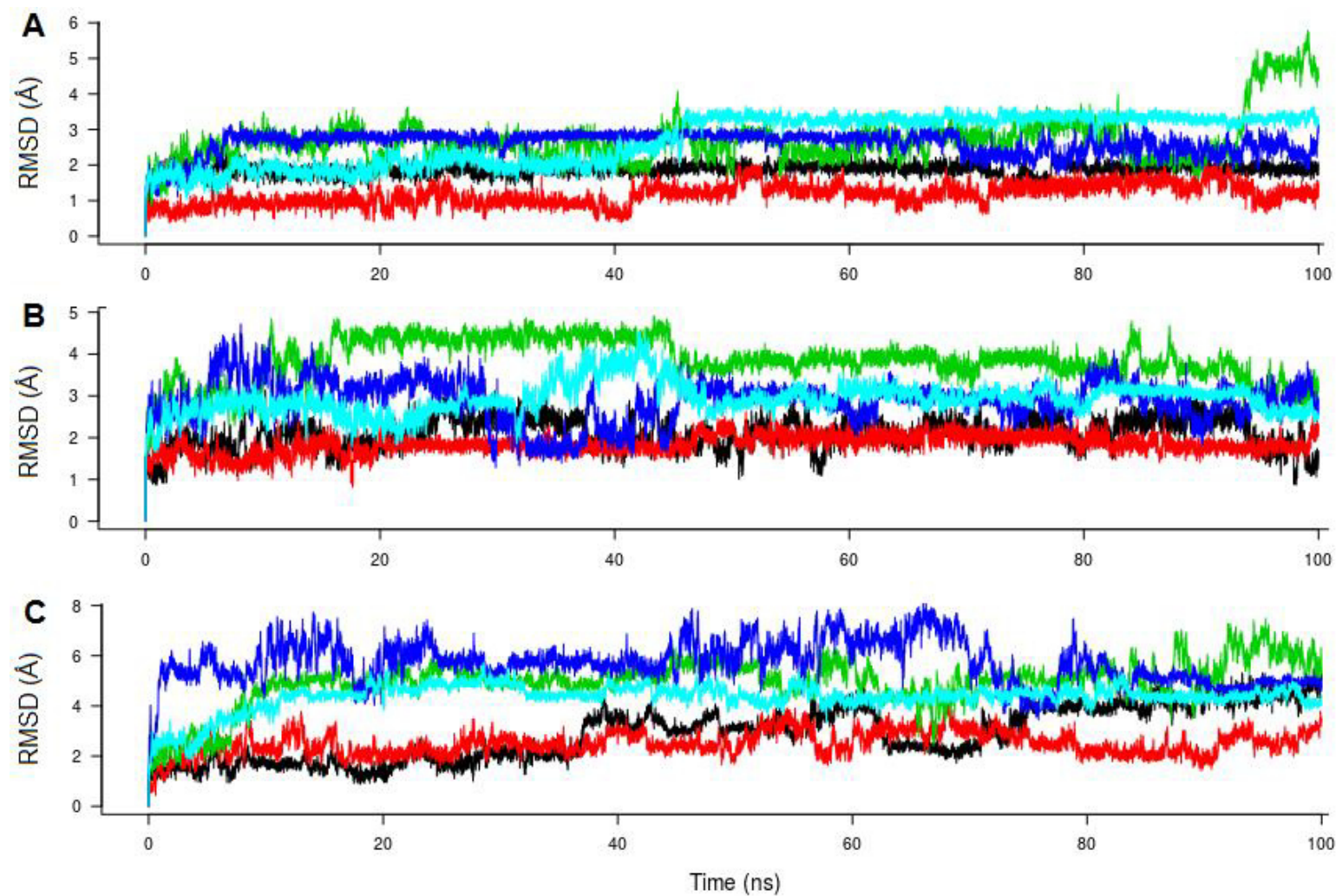
Supporting Figure 2. Determination of peptide stability with RMSD (A), RMSF (B) and radius of gyration (C). All plots of the five μ -conotoxins are in alphabetic order: (a) GIIIA, (b) KIIIA, (c) PIIIA, (d) SIIIA, and (e) SmIIIA. In general, the peptide native fold is represented in black, the peptide with one bridge opened is shown in red, and the peptide with two open bonds is in green. Values are given in Table S3.



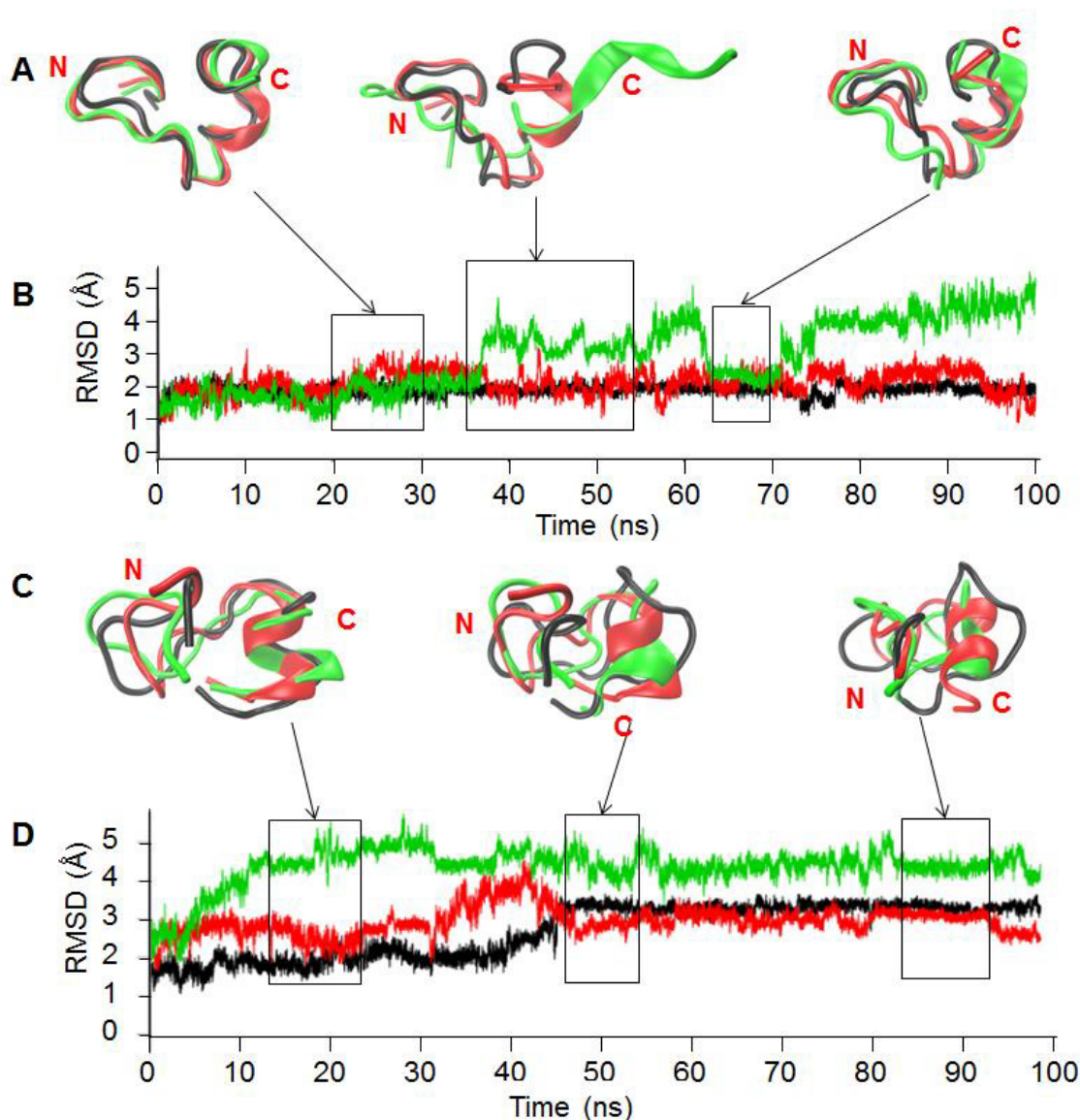
Supporting Figure 3. 1 μ s test trajectory of native μ -PIIIA grouped into different clusters based on the RMSD values.



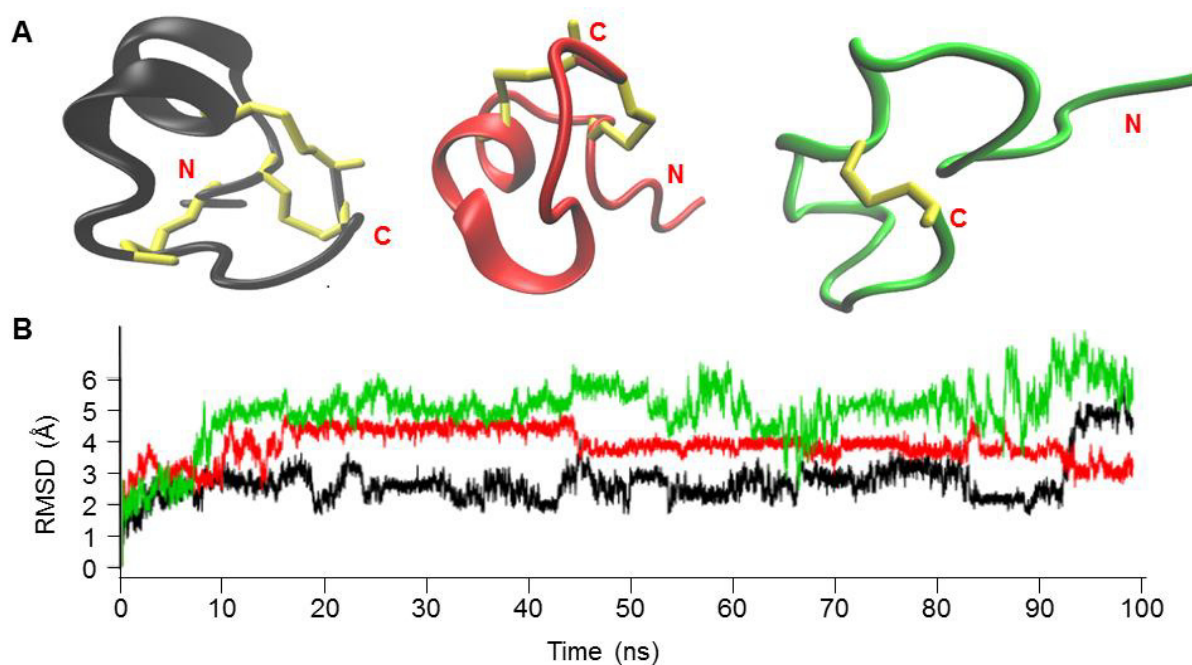
Supporting Figure 4. Backbone stability of the five μ -conotoxins (black GIIIA, red KIIIA, green PIIIA, blue SIIIA, cyan SmIIIA) in the native state (A), with disulfide bridge C2 – C5 removed (B) and with two disulfide bridges C2 – C5, C3 – C6 removed (C). In the case of μ -KIIIA C2-C4 and C3-C6 bonds were removed for the native peptide.



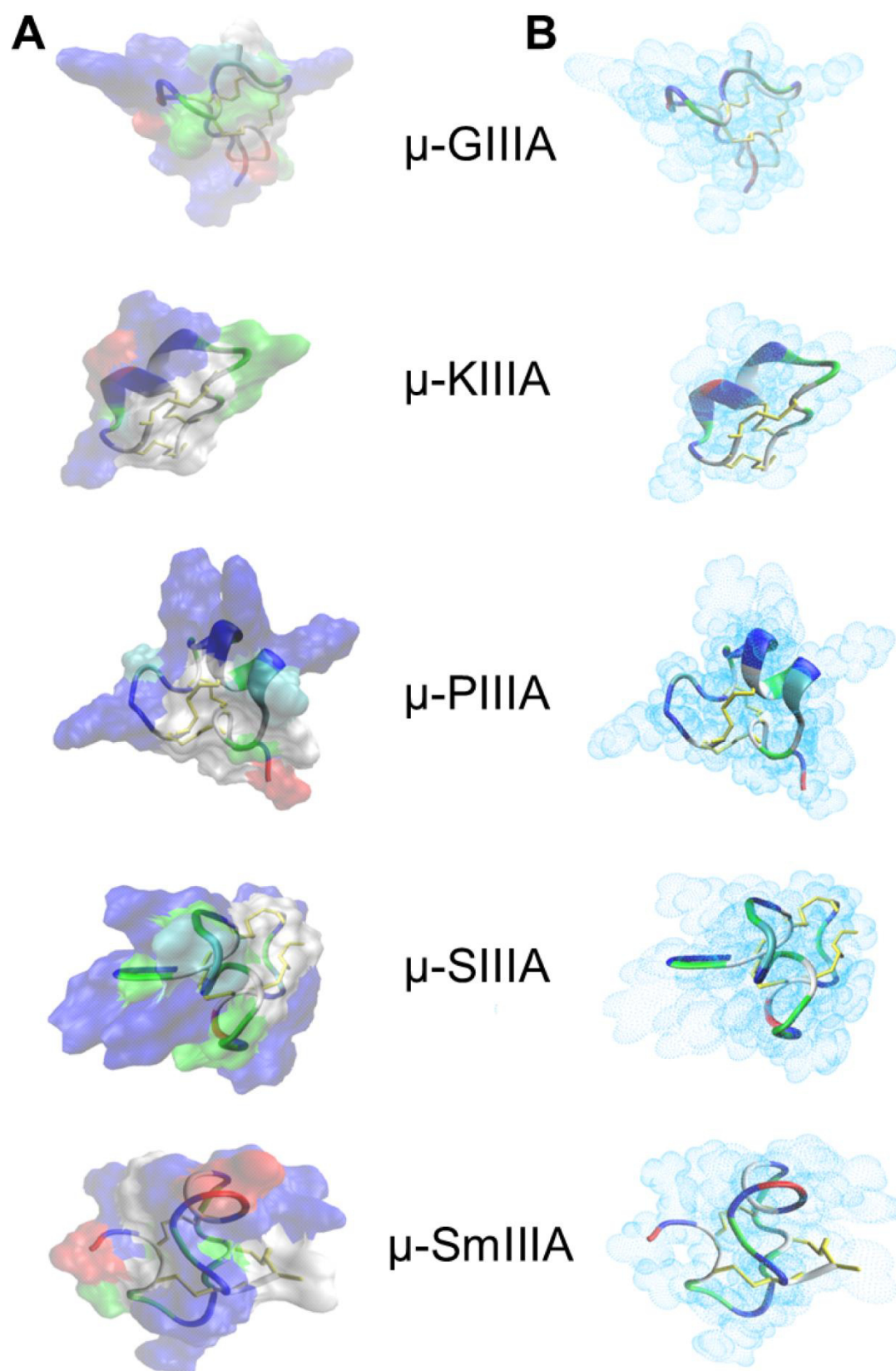
Supporting Figure 5. Backbone stability of μ -GIIIA and μ -SmIIIA. In general, the peptide native fold is represented in black, the peptide with one bridge opened is shown in red, and the peptide with two open bonds is in green. (A) μ -GIIIA average structures obtained and superimposed from 20 ns – 30 ns, 35 ns – 55 ns, and 65 ns – 70 ns. (B) RMSD for a simulation time of 100 ns of the respective peptide. In both cases, when compared with the native structure, the structure containing two disulfide bonds shows consistent structural stability. (C) μ -SmIIIA average structures obtained and superimposed from 15 ns – 25 ns, 45 ns – 55 ns, and 85 ns – 95 ns. (D) RMSD for a simulation time of 100 ns of the respective peptide. In contrast to the native structure the appearance of helices in the structures after disulfide bonds removed is noteworthy resulting in comparable stability. Helix formation in μ -GIIIA occurs between residues D12 and Q18. In μ -SmIIIA helix formation is between the residues W14 and R20.



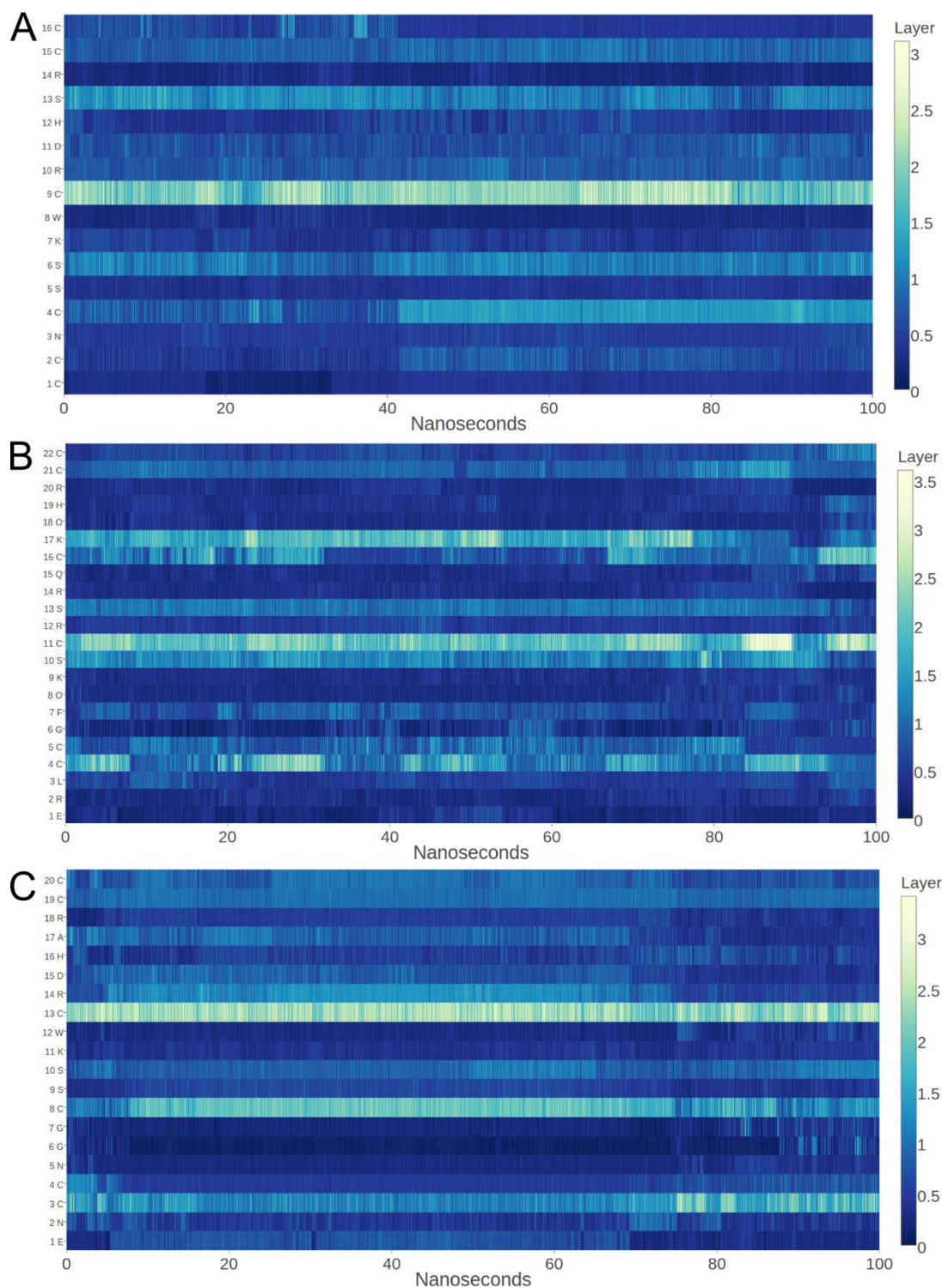
Supporting Figure 6. Representation of μ -PIIIA for native (black), one disulfide bridge (C2-C5, red) and two disulfide bridges (C2-C5, C3-C6, green) removed (A). Backbone RMSD for the three peptides for 100 ns (B).



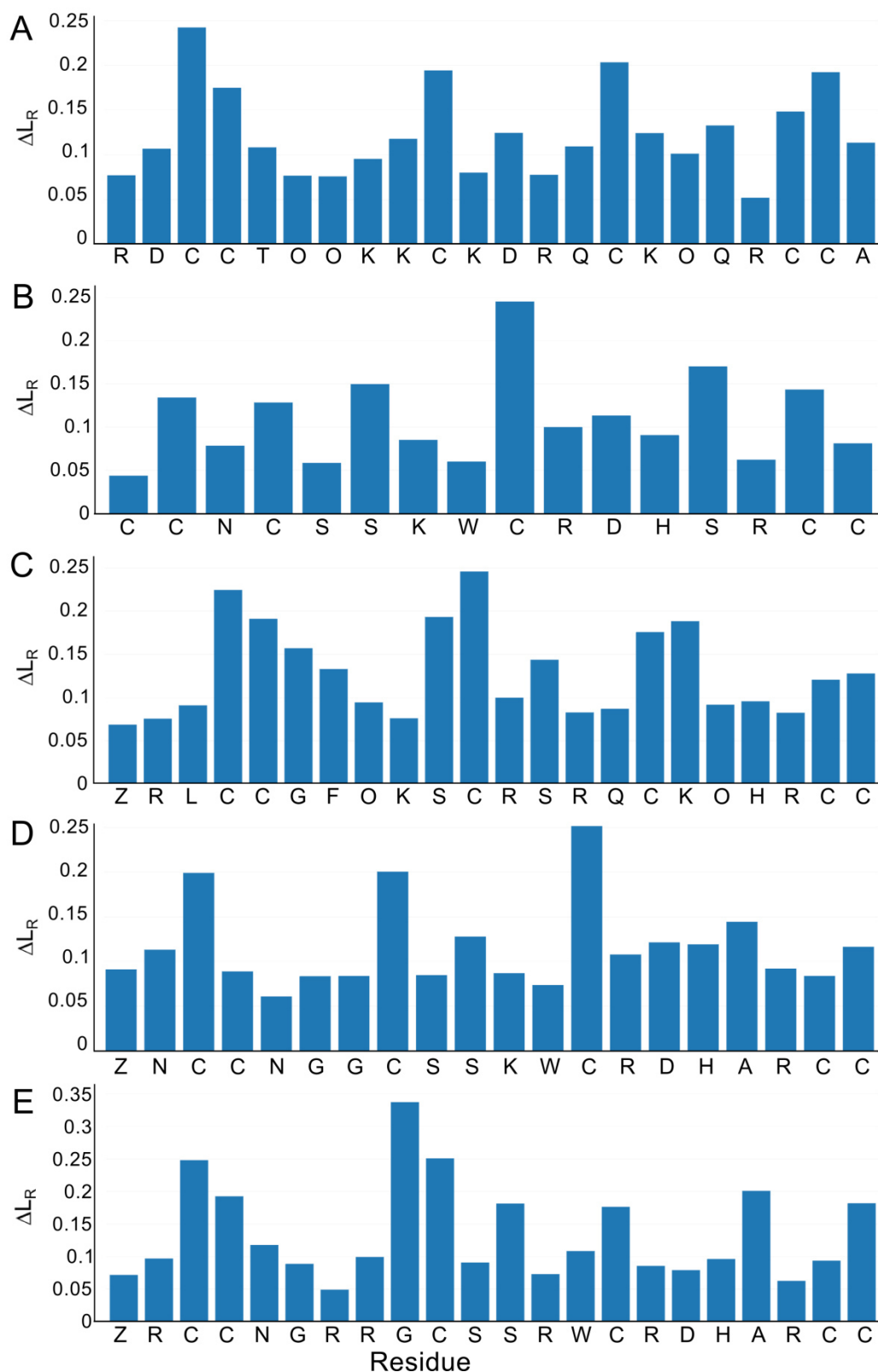
Supporting Figure 7. (A) Surface representation VDW and (B) Surface representation SAS (Solvent accessible surface) of the five μ -conotoxins.



Supporting Figure 8. The heatmaps show L_R for the residues of μ -conotoxins KIIIA (A), PIIIA (B) and SIIIA (C) over a 100 ns simulation. L_R values for each amino acid are stacked along the y-axis. L_R is mapped to the color scale. Dark blue refers to the surface of the molecule, since the outermost layer is assigned the value zero. Higher layer values, mapped to yellow, depict the amino acids that mostly form the core of the respective conotoxin. The plot quickly reveals the distribution of residues among the layers. Some of the residues clearly change their surface proximity throughout the simulation.



Supporting Figure 9. Bar chart showing ΔL_R for the residues of all five μ -conotoxins GIIIA (A), KIHA (B), PIHA (C), SIHA (D) and SmIIHA (E) for a 100 ns simulation. Larger bars show, that the corresponding residue changes its layer-membership frequently during the simulation. Thus, the plot quickly reveals the residues that compete with others, or these, that are stable with respect to their surface proximity.



Supporting Table 1. Comparison of sequence characteristics of the μ -conotoxins investigated in this study. Residues are highlighted according to their character: basic (blue), acidic (red), polar uncharged (green) and cysteine (yellow). (Z: pyroglutamate, O: 4-hydroxyproline) *All peptides were used as amides. Native μ -SmIIIA occurs as C-terminal acid, however, is usually used as amide.^[8,9] We used μ -SmIIIA as amide for reasons of comparison since both structures were found to be identical.^[9] #In general IC₅₀ values are given for expression in *Xenopus* oocytes with exception of SmIIIA, here only K_D is available.

μ -Conotoxin	Sequence	Net Charge	Structure Code (pdb)/ Structure Card ^[1]	Bioactivity on Na _v 1.4 [#]
GIIIA	-R ^{blue} D ^{red} CC ^{yellow} TOO ^{green} KK ^{yellow} C ^{yellow} K ^{blue} D ^{blue} R ^{blue} Q ^{yellow} C ^{yellow} KOQ ^{green} -R ^{blue} CC ^{yellow} A*	+5	S00082 (1TCG) ^[2]	19 +/- 1 nM ^[3]
KIIIA	---CC ^{yellow} N ^{green} ---C ^{yellow} SS ^{green} KK ^{yellow} W ^{yellow} CR ^{red} D ^{blue} H ^{blue} SR ^{blue} CC ^{yellow} -*	+3	S00129 (2LXG) ^[4]	90 +/- 17 nM ^[5]
PIIIA	ZRL ^{green} CC ^{yellow} GFO ^{green} KS ^{green} C ^{yellow} RS ^{green} R ^{blue} Q ^{yellow} C ^{yellow} KOH ^{green} -R ^{blue} CC ^{yellow} -*	+6	S00159 ^[6]	36 +/- 8 nM ^[3]
SIIIA	-ZN ^{green} CC ^{yellow} NG ^{green} --G ^{green} C ^{yellow} SS ^{green} KK ^{yellow} W ^{yellow} CR ^{red} D ^{blue} H ^{blue} AR ^{blue} CC ^{yellow} -*	+2	S00125 (BMRB 20025) ^[7]	130 +/- 15 nM ^[7]
SmIIIA	-ZR ^{blue} CC ^{yellow} NG ^{green} RR ^{blue} G ^{green} C ^{yellow} SS ^{green} RW ^{yellow} CR ^{red} D ^{blue} H ^{blue} SR ^{blue} CC ^{yellow} -*	+5	S00077 (1Q2J) ^[8]	0.22 +/- 1 nM ^[3]

Supporting Table 2. Comparison of disulfide loop length of the five μ -conotoxins. For the μ -conotoxins GIIIA, PIIIA, SIIIA and SmIIIA, the values corresponding to disulfide loops 1, 2 and 3 in the table is the number of residues present between the C1-C4, C2-C5 and C3-C6 disulfide bridges respectively. For μ -KIIIA, the disulfide loops 1, 2 and 3 correspond to the number of residues occurring between the disulphide bridges C1-C5, C2-C4 and C3-C6.

μ -Conotoxin	Disulfide loop 1	Disulfide loop 2	Disulfide loop 3
GIIA	11	15	10
KIIIA	15	6	11
PIIA	11	15	10
SIIIA	9	14	11
SmIIIA	11	16	11

Supporting Table 3. Assessment of stability using RMSD and radius of gyration for 100 ns MD simulation.

μ -Conotoxin	No. of Disulfide Bonds	Backbone RMSD (Å)	Radius of Gyration (Å)
GIIIA	3	1.8	7.6
	2	2.1	8.0
	1	2.8	8.2
KIIIA	3	1.1	6.7
	2	1.8	6.5
	1	2.5	7.1
PIIIA	3	2.3	8.0
	2	3.8	8.1
	1	4.9	8.5
SIIIA	3	2.6	7.0
	2	2.8	7.35
	1	5.7	8.5
SmIIIA	3	2.7	7.7
	2	2.9	7.6
	1	4.3	7.8

Supporting Table 4. Impact of the C2-C5 and C3-C6 disulfide bridge removal on the backbone and the functionally stable residues. Column 2 lists residues in each μ -conotoxin that are significant for binding activity as reported by Akondi *et al.*^[15] Column 3 of the table shows the difference (within brackets) in residue fluctuation between the native bridged form and the C2-C5 disulfide removed form of the five μ -conotoxins. Column 4 reports the difference in residue fluctuation between the native bridged form and the C2-C5 and C3-C6 disulfide removed conformations. The RMSF values are used as basis for determining increase or decrease of flexibility. A difference of close to 2 Å between RMSF values is considered significant and is marked in bold. *In the case of μ -KIIIA the C2-C4 disulfide bridge is removed. **Besides the general statement for the conserved basic residue between C3 and C4 essential for biological activity of μ -conotoxins, no order of basic residues is described so far.

μ -Conotoxin	Functional basic residues in order of significance on biological activity on Na _v 1.4	Comparison of flexibility on removal of the C2-C5* disulfide bridge (in Å)	Comparison of flexibility on removal of the C2-C5* and C3-C6 bridge disulfide bridge (in Å)
GIIIA	R13 > R19 > R1, K8, K11, K16 > K9 ^[10,11]	R13 (0), R19 (0.5), R1 (0.7), K8(0.9), K11(0.1), K16 (0) K9 (0.1)	R13 (0.5), R19 (2.6) , R1 (0.8), K8(0.7), K11(0.5), K16 (2.8) , K9 (0)
KIIIA	H12 > R14 > K7 > R10 ^[12,13]	H12 (0.2), R14 (0.5), K7 (0.4), R10 (0.3)	H12 (0.8), R14 (0.9), K7 (0.5), R10 (0.4)
PIIIA	R14 > K17 > R12, R20, R2, K9 ^[14]	R14 (1.3), K17 (0.7), R12 (0), R20 (0.2), R2 (0.9), K9 (1)	R14 (1.4), K17 (0.2), R12 (0.6), R20 (1.4), R2 (0.9), K9 (0.3)
SIIIA	R14 > K11, H16, R18**	R14 (0.4), K11 (0.3), H16 (0.4), R18 (1.1)	R14 (1.7) , K11 (1.2), H16 (1.9) , R18 (2.8)
SmIIIA	R13 > R7, R8, R16, R20**	R13 (1.4), R7 (1.6) , R8 (1.2), R16 (0.6), R20 (0.1)	R13 (1.8) , R7 (1.9) , R8 (2.2Å) , R16 (0.6), R20 (0.1)

Supporting References

- [1] Q. Kaas, R. Yu, A. H. Jin, S. Dutertre, D. J. Craik, *Nucleic Acids Res.* **2012**, *40*, 325–330.
- [2] K. Wakamatsu, D. Kohda, H. Hatanaka, J. M. Lancelin, Y. Ishida, M. Oya, H. Nakamura, F. Inagaki, K. Sato, *Biochemistry* **1992**, *31*, 12577–12584.
- [3] M. J. Wilson, D. Yoshikami, L. Azam, J. Gajewiak, B. M. Olivera, G. Bulaj, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10302–10307.
- [4] K. K. Khoo, K. Gupta, B. R. Green, M. M. Zhang, M. Watkins, B. M. Olivera, P. Balaram, D. Yoshikami, G. Bulaj, R. S. Norton, *Biochemistry* **2012**, *51*, 9826–9835.
- [5] M. M. Zhang, B. R. Green, P. Catlin, B. Fiedler, L. Azam, A. Chadwick, H. Terlau, J. R. McArthur, R. J. French, J. Gulyas, et al., *J. Biol. Chem.* **2007**, *282*, 30699–30706.
- [6] A. A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Köhl, A. Mischo, G. Buntkowsky, M. Görlach, S. H. Heinemann, D. Imhof, *Angew. Chem. Int. Ed.* **2012**, *51*, 4058–4061.
- [7] S. Yao, M. M. Zhang, D. Yoshikami, L. Azam, B. M. Olivera, G. Bulaj, R. S. Norton, *Biochemistry* **2008**, *47*, 10940–10949.
- [8] D. W. Keizer, P. J. West, E. F. Lee, D. Yoshikami, B. M. Olivera, G. Bulaj, R. S. Norton, *J. Biol. Chem.* **2003**, *278*, 46805–46813.
- [9] E. Fuller, B. R. Green, P. Catlin, O. Buczek, J. S. Nielsen, B. M. Olivera, G. Bulaj, *Febs J.* **2005**, *272*, 1727–1738.
- [10] G. Choudhary, M. P. Aliste, D. P. Tieleman, R. J. French, S. C. Dudley Jr., *Channels* **2007**, *1*, 344–352.
- [11] K. Sato, Y. Ishida, K. Wakamatsu, R. KJato, H. Honda, Y. Ohizumi, H. Nakamura, M. Ohya, J.-M. Lancelin, D. Kohda, et al., *J. Biol. Chem.* **1991**, *266*, 16989–16991.
- [12] J. R. McArthur, G. Singh, D. McMaster, R. Winkfein, D. P. Tieleman, R. J. French, *Mol. Pharmacol.* **2011**, *80*, 573–84.
- [13] S. Mahdavi, S. Kuyucak, *Toxins* **2014**, *6*, 3454–3470.
- [14] J. R. McArthur, G. Singh, M. L. O'Mara, D. McMaster, V. Ostroumov, D. P. Tieleman, R. J. French, *Mol. Pharmacol.* **2011**, *80*, 219–227.
- [15] K. B. Akondi, M. Muttenthaler, Q. Kaas, D. J. Craik, R. J. Lewis, P. F. Alewood, *Chem. Rev.* **2014**, *114*, 5815–5847.

Supplementary Information

Conformational μ -PIIIA peptide isomers revisited: The impact of disulfide connectivity on structure and bioactivity

Pascal Heimer,^[a] Alesia A. Tietze,^[b] Enrico Leipold,^[c] Desiree Kaufmann,^[d] Charlotte A. Bäuml,^[a] Ajay A. Paul George,^[a] Astrid Maass,^[e] Jan Hamaekers,^[e] Anja Resemann,^[f] Franz J. Mayer,^[f] Detlev Suckau,^[f] Oliver Ohlenschläger,^[g] Stefan H. Heinemann,^[c] Daniel Tietze^{[d]*} and Diana Imhof^{[a]*}

Supporting Information

Supporting Text

Synthesis Notes. The combination of the standard acetamidomethyl (Acm), *tert*-butyl (tBu) and trityl (Trt) cysteine protecting groups was used for the synthesis of μ -PIIIA isomers (Figure S2, Table S1).^{1–3} Partially Cys-protected linear precursors were produced and subjected to stepwise oxidation in the following order regarding side chain deprotection: (1) oxidation of formerly Trt-protected Cys-pair, (2) deprotection and oxidation of the Acm-protected Cys-pair, and (3) deprotection and oxidation of the tBu-protected Cys-pair. Reaction control was carried out after each oxidation step to monitor the different behavior of the isomers during folding. Under the conditions applied, several isomers (**1–3**, **6**, **14**, **15**) were produced as main products in higher yields compared to other peptides (**7**, **8**, **10**, **12**, **13**), which contained side products (Figure S2, Table S2). In addition, for isomers **4**, **5**, **9**, and **10**, two main products were obtained. These findings were rather unexpected with respect to the chemical synthesis strategy applied because critical aspects such as high temperature loss of the Acm group⁴ were considered in advance and, furthermore, an already established protecting group strategy was used.¹ However, obvious side reactions take place at the final deprotection and oxidation step, where in some cases two completely oxidized species with different retention times were observed. This can only be explained by disulfide scrambling during oxidation with diphenylsulfoxide and trimethylsilylchloride, described by Szabo *et al.*³ In this report, scrambling between the former Acm- and tBu-protected cysteines of the α -conotoxin GI occurred during the synthesis process. However, this is in contrast to other protocols in which such a scrambling was not observed.^{5,6}

HPLC co-elution of isomer mixtures. In order to understand our previous findings obtained from an oxidative self-folding approach⁷, co-elution of different mixtures of isomers was performed (Figure S6). In general, equal amounts of three isomers in each case were mixed together and directly injected in the respective HPLC system equipped with either a C4 or a C18 column. In case of the RP18 column separations, only for one mixture (isomers **1**, **2**, **3**) three peaks were detectable in the elution profile, yet not baseline separated (Figure S6a). Further mixtures of relevant isomers sharing a distinct disulfide bond such as **1**, **3**, and **15** (C11-C16) or **3**, **6**, and **9** (C21-C22) (Figure S6b, c, e, f) resulted in two peaks. One exception was the mixture of isomers **3**, **4**, and **5** sharing bridge C4-C5 (Figure S6d), where only one peak was observed. In our opinion, this could be the reason why previous NMR experiments with isomer **3** failed due to contamination with other isomers that were obviously not present or as dominant for isomers **1** and **2** in the former study.⁷ Since the measurements using the C18 column revealed no significant difference in the retention times

of the isomer compared to those performed on a C8 column, the co-elution experiment was not carried out using C8 material. Instead, another HPLC instrument equipped with a C4 column was chosen (Figure S6 g-l). However, observations were similar compared to the results obtained with the C18 column. One exception was the mixture of isomers **1**, **2**, and **3**, where only two peaks were identified instead of three as found on the C18 column. These findings clearly indicate that resolution of μ -PIIIA isomers is not possible with the chromatographic methods commonly used to separate peptides and small proteins. Consequently, co-elution does not provide sufficient proof of identity for disulfide-rich peptides and proteins from biological, synthetic or recombinantly expressed material if not demonstrated for the respective standard molecules in suitable mixtures.

NMR structure determination. For some μ -PIIIA isomers (with the exception of **1**, **2**, **9**, **11**, **14**, **15**) we observed an additional set of NMR cross peaks due to an intermediate or high backbone deviation. With respect to the MS/MS data, which did not indicate any major impurities resulting from alternative isomers (except for **6** and **12**), the additional NMR signals probably resulted from other toxin conformations being in slow exchange with the major conformation. Similar observations were reported for the three possible α -GI isomers possessing two disulfide bonds, yielding a very rigid backbone conformation for the natively folded isomer and flexible conformations for the non-native versions.⁸ The same was recently published for α -conotoxin Pu1.2,⁹ in which the C1-C2, C3-C4 isomer ensemble displayed the highest RMSD value and the lowest structural definition.

Distance constraints derived from NOE data do not necessarily lead to a unique structure thus being compatible with multiple disulfide bond patterns.¹⁰ Even inter-cysteine NOEs, which are typically recognized as indicative for unique disulfide connectivity can lead to biased results for closely packed disulfide networks. Poppe and coworkers found that the disulfide pattern is underdetermined for approximately 30% of all NMR structures deposited in the PDB database.¹⁰ In order to test the NMR restraints with respect to the ambiguity of predicting the disulfide pattern, two alternative structure calculations based on the individual sets of NMR restraints were performed for all 15 isomers. These calculations revealed plausible alternative structures for isomers **1**, **2**, **10**, and **14**, and to a lesser extent also for **4**, **5**, and **8** (Table S5). This clearly indicates the importance of an unambiguous assignment of the disulfide bond patterns by additional methods, such as the aforementioned MS/MS analysis.

However, re-evaluation of the results from the self-folding approach described earlier for isomers **1-3**⁷ revealed contamination with one or more isomers. This was supported by comparing our previous 2D TOCSY NMR spectra with the data obtained herein confirming that the former isomer **2** also contained **1** and vice versa. In addition, the former isomer **3**

contained significant amounts of other isomers, most likely isomers **4** and **6**, thus rationalizing the poor quality of the NOE data at that time.⁷

Electrophysiological experiments. Human SCN4A (encoding the Na_v1.4 channel α subunit, UniProt ID P35499) on a plasmid with CMV promoter was transiently expressed in HEK 293 cells as shown previously.¹¹ Co-transfection of a plasmid encoding the CD8 antigen ensured visual detection of transfected cells with CD8-specific Dynabeads (Deutsche Dynal, Hamburg, Germany). Currents were measured with the whole-cell patch clamp method 24–48 h after transfection.¹¹ The patch pipettes contained (in mM): 35 NaCl, 105 CsF, 10 EGTA (ethylene glycol bis(2-amino-ethylether)tetraacetic acid), 10 HEPES (pH 7.4 with CsOH). The bath solution contained (in mM): 150 NaCl, 2 KCl, 1.5 CaCl₂, 1 MgCl₂, 10 HEPES (pH 7.4 with NaOH). Holding potential was –120 mV, Na⁺ currents were elicited with depolarizing steps to –20 mV. Series resistance was corrected electronically up to 80%. Peptides, diluted in the bath solution, were applied focally to cells under consideration with a fine-tipped glass capillary. Time course of peak current decrease after peptide application was described with single-exponential functions.

Homology modeling. High-affinity μ -conotoxin binding to the pore requires optimal positioning of the residues that contribute to the toxin receptor. However, the orientation of key residues responsible for high affinity toxin binding mainly depends on the global fold of the toxin. In case of cysteine rich toxins such as μ -PIIIA, the disulfide bond pattern most likely determines the global fold. Therefore, the question arises how do individual disulfide bonds contribute determine the global fold of the toxin and how does this affect the orientation of key residues. To answer these questions, we aimed at using the aforementioned NMR structures and a homology model of Na_v1.4 for molecular docking studies.

Although crystal structures of voltage-gated Na⁺ channels (Na_vRh PDB ID: 4DXW¹², Na_vAB-Ile217Cys, -wt PDB ID: 3RVY¹³, 4EKW¹⁴) are available in the RCSB database, including the recently solved structure of an eukaryotic Nav channel (Na_vPaS, pdb ID 5X0M)¹⁵ the structure of the mammalian Na_v1.4 channel is not yet described. Therefore, we built an extended pore model including essential parts of the loop structures that connect the transmembrane segments S5 and S6 (see also Figure S9) of each pore domain with the pore helices SS1 and SS2 and the selectivity filter. So far, the recently determined crystal structure of a voltage-gated Ca²⁺ channel (PDB ID 3jbr)¹⁶, which also exhibits extended loop structures compared to Na_v1.4 (Figure S9) being partly similar in length and sequence (sequence identity 21 %) according to the conservation score (Figure S9), was identified as the only available template for the anticipated loop modeling approach at that time. The structure of Na_vPaS was not available, when this study was started, but is in good agreement

with our modelled structure (r.m.s.d. 1.9 Å, Figure S10). However, loops connecting S5 and SS1 of domain I and domain III were only partly modeled due to their much smaller size (domain I) or their poor sequence similarity (domain III, see Figure S9). The remaining loops were either correct in length or were modified by inserting or deleting the respective residues accordingly (Figure S9). Furthermore, due to the high structural similarity of the pore modules of Ca_v1.1 and Na_vAB (PDB ID 3rvy, r.m.s.d. 1.56 Å, sequence identity 26.2 %) comprising transmembrane segments S5 and S6 and pore helices SS1 and SS2 of each pore domain, Ca_v1.1 was also considered as a suitable template for the pore module of Na_v1.4. In order to retain the template structure as much as possible, the template sequence was only changed according to the target sequence for those regions that are essential for toxin binding (Figure S9). The crude homology model was then re-aligned with the pore module of the crystal structure of Na_vAB re-adjusting backbone coordinates and side chain orientations accordingly and followed by a 2-ns refinement simulation in a membrane environment.

Supporting Table 1. Protecting group strategy used for the preparation of μ -PIIIA isomers.

Isomer	Connectivity	Trt*	Acm*	tBu*
1**	C4-C21, C5-C22, C11-C16	C4-C21	C5-C22	C11-C16
2**	C4-C16, C5-C21, C11-C22	C4-C16	C11-C22	C5-C21
3**	C4-C5, C11-C16, C21-C22	C4-C5	C11-C16	C21-C22
4	C4-C5, C11-C21, C16-C22	C4-C5	C11-C21	C16-C22
5	C4-C5, C11-C22, C16-C21	C4-C5	C11-C22	C16-C21
6	C4-C11, C5-C16, C21-C22	C21-C22	C5-C16	C4-C11
7	C4-C11, C5-C21, C16-C22	C4-C11	C5-C21	C16-C22
8	C4-C11, C5-C22, C16-C21	C16-C21	C5-C22	C4-C11
9	C4-C16, C5-C11, C21-C22	C21-C22	C4-C16	C5-C11
10	C4-C16, C5-C22, C11-C21	C11-C21	C5-C22	C4-C16
11	C4-C21, C5-C11, C16-C22	C5-C11	C4-C21	C16-C22
12	C4-C21, C5-C16, C11-C22	C5-C16	C11-C22	C4-C21
13	C4-C22, C5-C11, C16-C21	C4-C22	C16-C21	C5-C11
14	C4-C22, C5-C16, C11-C21	C4-C22	C11-C21	C5-C16
15	C4-C22, C5-C21, C11-C16	C4-C22	C5-C21	C11-C16

*Protecting group used for the given cysteine pair.

**These isomers have been described previously⁷.

Supporting Table 2. Analytical characterization of μ -PIIIA isomers produced in this study.

Isomer	t_R (C18) [min]	t_R (C8) [min]	MW (calc.) [mono]	MW (found) [M+H] ⁺	NMR (MHz)
1	16.4	17.7	2603.1	2604.2	600
2	16.2	17.1	2603.1	2604.2	600
3	17.6	18.8	2603.1	2604.2	700
4	18.5	19.6	2603.1	2604.1	700
5	17.5	18.7	2603.1	2604.1	600
6	17.4	18.3	2603.1	2604.1	700
7	1: 16.9* 2: 18.1	1: 17.0 2: 18.4	2603.1	2604.1	700
8	17.4	18.5	2603.1	2604.2	700
9	1: 16.1* 2: 19.6	1: 17.1 2: 20.6	2603.1	2604.1	700
10	17.4	18.6	2603.1	2604.3	600
11	15.6	16.8	2603.1	2604.2	900
12	17.4	18.5	2603.1	2604.1	700
13	19.9	20.9	2603.1	2604.2	700
14	16.3	17.5	2603.1	2604.1	700
15	16.4	17.3	2603.1	2604.0	900

t_R , retention time; MW, molecular weight; calc., calculated.

*Isomers occurring as two or more peaks in HPLC analysis were separately analyzed, and the respective fraction used for NMR structure determination is marked.

Supporting Table 3. Results from MALDI MS/MS sequencing of partially reduced and alkylated μ -PIIIA species of the respective isomers as derived from internal fragments and BioTools analysis (Bruker). The carbamidomethylated species were analyzed to obtain the corresponding results. The disulfide connectivity obtained in MS/MS, which correlates to the envisaged synthesis product, is marked in bold. Isomers, in which the expected bridging is not correlating with the MS data are marked in italics. Retention times (first column) are given relative to the 6x carbamidomethylated version. The unmodified (0 CAM) and 6x carbamidomethylated versions cannot be used for structure elucidation and are therefore marked with “-”.

Isomer 1 (C4-C21, C5-C22, C11-C16)					
$t_R^{[a]}$ (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
-0.5	2724.2	2	1497	ZRLC*CGFOKSCRSRQCKOHRCC*-NH ₂	C4–C21
	2838.3	4	1554	ZRLCC*GFOKSC*RSRQC*KOHRCC*-NH ₂	C4–C21
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
2.0	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
	2724.2	2	1440 + 1497	ZRLC'C*GFOKSCRSRQCKOHRCC*-NH ₂	C5–C22 and C4–C21
Disulfide connectivity with highest score: C4–C21, C5–C22, C11–C16					
Isomer 2 (C4-C16, C5-C21, C11-C22)					
$t_R^{[a]}$ (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
0.5	2724.2	2	1497	ZRLC*CGFOKSCRSRQCKOHRCC-NH ₂	C4–C16
	2838.3	4	1497 + 1554	ZRLC*CGFOKSC*RSRQC*KOHRCC*-NH ₂	C5–C21
1.4	2724.2	2	1440 + 1497	ZRL(CC)*GFOKSCRSRQCKOHR(CC)*-NH ₂	<i>C4–C22 and C5–C21</i>

	2838.3	4	1497	ZRLC*C*GFOKSCRSRQCKOHR C*C*-NH ₂	C4-C5, C4-C22, C5-C22
2.4	2724.2	2	1440	ZRLC*C*GFOKSCRSRQCKOHRCC-NH ₂	C4-C5, C4-C22, C5-C22
2.8	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4-C16, C5-C21, C11-C22					
Isomer 2 (C4-C16, C5-C21, C11-C22) from oxidative self-folding approach in buffer ⁷					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
-1.0	2724.2	2	1497	ZRLCC*GFOKSCRSRQCKOHR C*C*-NH ₂	C5-C21
	2838.2	4	1554	ZRLC*C*GFOKSCRSRQC*KOHR C*C*-NH ₂	C11-C22
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHR C*C*-NH ₂	-
0.2	2724.2	2	1497	ZRLC*CGFOKSCRSRQC*KOHRCC-NH ₂	C4-C16
	2838.2	4	1554	ZRLC*C*GFOKSCRSRQC*KOHR C*C*-NH ₂	C11-C22
1.1	2724.2	2	1440 + 1497	ZRL(CC)*GFOKSCRSRQCKOHR(CC)*-NH ₂	C5-C22 and C5-C21
	2838.5	4	1497 + 1611	ZRLCCGFOKSC*RSRQC*KOHR C*C*-NH ₂	C5-C22
2.8	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4-C16, C5-C21, C11-C22 , <i>minor C4-C21, C5-C22, C11-C16</i>					
Isomer 3 (C4-C5, C11-C16, C21-C22)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHR C*C*-NH ₂	-
1.4	2724.2	2	1440	ZRLC*C*GFOKSCRSRQCKOHRCC-NH ₂	C4-C5

	2838.3	4	1497	ZRLC*C*GFOKSCRSRQCKOHRCC*-NH ₂	C11–C16
3.1	2724.2	2	1554	ZRLCCGFOKSC*RSRQC*KOHRCC-NH ₂	C11–C16
	2838.3	4	1611	ZRLCCGFOKSC*RSRQC*KOHRCC*-NH ₂	C4–C5
4.1	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4–C5, C11–C16, C21–C22					
Isomer 4 (C4-C5, C11-C21, C16-C22)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
0.3	2724.4	2	1440	ZRLC*C*GFOKSCRSRQCKOHRCC-NH ₂	C4–C5
	2838.5	4	1497	ZRLC*C*GFOKSCRSRQC*KOHRCC*-NH ₂	C11–C21
2.7	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
3.5	2724.4	2	1497	ZRLCCGFOKSC*RSRQC*KOHRCC-NH ₂	C16–C22
	2838.5	4	1611	ZRLCCGFOKSC*RSRQC*KOHRCC*-NH ₂	C4–C5
Disulfide connectivity with highest score: C4–C5, C11–C21, C16–C22					
Isomer 5 (C4-C5, C11-C22, C16-C21)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
1.4	2724.2	2	1497	ZRLCCGFOKSC*RSRQCKOHRCC*-NH ₂	C11–C22
3.1	2724.2	2	1554	ZRLCCGFOKSCRSRQC*KOHRCC*-NH ₂	C16–C21

	2838.3	4	1611	ZRLCCGFOKSC*RSRQC*KOHRC*C*-NH ₂	C4–C5
4.1	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
	2724.2	2	1440	ZRLC*C*GFOKSCRSRQCKOHRCC-NH ₂	C4–C5
	2838.3	4	1497	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂	C16–C21
Disulfide connectivity with highest score: C4–C5, C11–C22, C16–C21					
Isomer 6 (C4-C11, C5-C16, C21-C22)					
$t_R^{[a]}$ (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRC*C*-NH ₂	-
3.1	2724.4	2	1554 + 1497	ZRLCCGFOKSC'RSRQC'KOHRC*C*-NH ₂	<i>C11–C16</i> and C21–C22
	2838.4	4	1611 + 1554	ZRLCCGFOKSC*RSRQC*KOHRC*C*-NH ₂ ZRLCC*GFOKSCRSRQC*KOHRC*C*-NH ₂	<i>C4–C5</i> and C4–C11
3.4	2724.4	2	1497	ZRLCC*GFOKSCRSRQC*KOHRCC-NH ₂	C5–C16
	2838.4	4	1611 + 1554	ZRLCCGFOKSC*RSRQC*KOHRC*C*-NH ₂ ZRLCC*GFOKSCRSRQC*KOHRC*C*-NH ₂	<i>C4–C5</i> and C4–C11
4.2	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4–C11, C5–C16, C21–C22 , <i>minor C4–C5, C11–C16, C21–C22</i>					
Isomer 7 (C4-C11, C5-C21, C16-C22)					
$t_R^{[a]}$ (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRC*C*-NH ₂	-
1.3	2724.2	2	1497	ZRLCCGFOKSC'RSRQC'KOHRC*C*-NH ₂	C5–C21

	2838.2	4	1611 +1554	ZRLC*CGFOKSC*RSRQC*KOHRC*C-NH ₂ ZRLCC*GFOKSCRSRQC*KOHRC*C*-NH ₂	C5-C22 and C4-C11
2.7	2724.2	2	1497	ZRLCC*GFOKSCRSRQCKOHRCC*-NH ₂	C5-C21
	2838.2	4	1497 +1554	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂ ZRLCC*GFOKSCRSRQC*KOHRC*C*-NH ₂	C16-C21 and C4-C11
3.7	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4-C11, C5-C21, C16-C22 , <i>minor C4-C11, C5-C22, C16-C21</i>					
Isomer 8 (C4-C11, C5-C22, C16-C21)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRC*C*-NH ₂	-
1.4	2724.2	2	1440	ZRLCC*GFOKSCRSRQCKOHRCC*-NH ₂	C5-C22
	2838.3	4	1497	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂	C16-C21
3.1	2724.2	2	1497 + 1554	ZRLC'CGFOKSC'RSRQC*KOHRC*C-NH ₂	C4-C11 and C16-C21
	2838.3	4	1611	ZRLC*CGFOKSC*RSRQC*KOHRC*C-NH ₂	C5-C22
4.0	2724.2	2	1497	ZRLC*CGFOKSC*RSRQCKOHRCC-NH ₂	C4-C11
	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4-C11, C5-C22, C16-C21					
Isomer 9 (C4-C16, C5-C11, C21-C22)					
Retention time ^[a] (min)	Mass found (m/z)	No. of CAM- modified residues (0-6)	Internal Fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRC*C*-NH ₂	-

1.4	2724.2	2	1440	ZRLCC*GFOKSCRSRQCKOHRCC*-NH ₂	C5-C11
	2838.3	4	1554	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂	C4-C16
3.1	2724.2	2	1497	ZRLC'CGFOKSC'RSRQC*KOHRCC*-NH ₂	C4-C16
	2838.3	4	1554	ZRLC*CGFOKSC*RSRQC*KOHRCC*-NH ₂	C21-C22
4.0	2724.2	2	1497	ZRLC*CGFOKSC*RSRQCKOHRCC-NH ₂	C21-C22
	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-

Disulfide connectivity with highest score: **C4-C16, C5-C11, C21-C22**

Isomer **10** (C4-C16, C5-C22, C11-C21)

$t_R^{[a]}$ (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
0.7	2724.4	2	1497 + 1440	ZRLC*CGFOKSCRSRQC*KOHRCC-NH ₂ ZRLCC*GFOKSCRSRQCKOHRCC*-NH ₂	C4-C16 and C5-C22
	2838.4	4	1497 + 1611	ZRLC*C*GFOKSCRSRQC*KOHRCC*-NH ₂ ZRLC*CGFOKSC*RSRQC*KOHRCC*-NH ₂	C11-C21 and C5-C22
2.9	2724.4	2	1497 + 1554	ZRLC*CGFOKSCRSRQC*KOHRCC-NH ₂ ZRLCCGFOKSC*RSRQCKOHRCC*-NH ₂	C4-C16 and C11-C21
	2838.5	4	1611 + 1554	ZRLC*CGFOKSC*RSRQC*KOHRCC*-NH ₂ ZRLCC*GFOKSC*RSRQCKOHRCC*-NH ₂	C5-C22 and C4-C16
3.6	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-

Disulfide connectivity with highest score: **C4-C16, C5-C22, C11-C21**

Isomer **11** (C4-C21, C5-C11, C16-C22)

$t_R^{[a]}$ (min)	MW found	No. of CAM- modified	Internal fragment	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
----------------------	-------------	-------------------------	----------------------	---------------------------------------	------------------------

	(m/z)	residues (0-6)	found ^[b] (m/z)		
-1.0	2724.3	2	1497	ZRLCC*GFOKSC*RSRQCKOHRCC-NH ₂	C5–C11
	2838.4	4	1554	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂	C16–C22
-0.2	2724.3	2	1497	ZRLC*CGFOKSCRSRQCKOHRCC*-NH ₂	C4–C21
	2838.4	4	1554	ZRLC*C*GFOKSC*RSRQCKOHRCC*-NH ₂	C16–C22
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
2.2	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4–C21, C5–C11, C16–C22					
Isomer 12 (C4-C21, C5-C16, C11-C22)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
1.2	2724.3	2	1440	ZRLC*CGFOKSCRSRQCKOHRCC*-NH ₂	C16–C22
	2838.4	4	1497	ZRLC*C*GFOKSCRSRQC*KOHRCC*-NH ₂	C4–C21
2.8	2610.1	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
3.1	2724.3	2	1497	ZRLCC*GFOKSCRSRQC*KOHRCC-NH ₂	C4–C21
	2838.4	4	1611	ZRLCC*GFOKSC*RSRQC*KOHRCC*-NH ₂	C16–C22
Disulfide connectivity with highest score: C4–C21, C5–C16, C11–C22					
Isomer 13 (C4-C22, C5-C11, C16-C21)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity

0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
1.1	2724.3	2	1440	ZRLC*CGFOKSCRSRQCKOHRCC*-NH ₂	C4–C22
	2838.4	4	1497	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	C16–C21
2.7	2724.2	2	1554	ZRLCCGFOKSCRSRQC*KOHRCC*-NH ₂	C16–C21
	2838.3	4	1611	ZRLCC*GFOKSC*RSRQC*KOHRCC*-NH ₂	C4–C22
3.6	2610.1	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4–C22, C5–C11, C16–C21					
Isomer 14 (C4-C22, C5-C16, C11-C21)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity
0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRCC*-NH ₂	-
1.3	2724.2	2	1440	ZRLC*CGFOKSCRSRQCKOHRCC*-NH ₂	C4–C22
	2838.3	4	1497	ZRLC*C*GFOKSCRSRQC*KOHRCC*-NH ₂	C11–C21
3.0	2724.2	2	1497 + 1554	ZRLCC*GFOKSC*RSRQC*KOHRCC*-NH ₂	C5–C16 and C11–C21
	2838.3	4	1611	ZRLCC*GFOKSC*RSRQC*KOHRCC*-NH ₂	C4–C22
3.9	2610.1	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
	2724.2	2	1497	ZRLCC*GFOKSCRSRQC*KOHRCC-NH ₂	C5–C16
Disulfide connectivity with highest score: C4–C22, C5–C16, C11–C21					
Isomer 15 (C4-C22, C5-C21, C11-C16)					
t _R ^[a] (min)	MW found (m/z)	No. of CAM- modified residues (0-6)	Internal fragment found ^[b] (m/z)	Sequence (*carbamidomethylation, CAM)	Disulfide connectivity

0	2952.0	6	1611	ZRLC*C*GFOKSC*RSRQC*KOHRC*C*-NH ₂	-
0.7	2724.2	2	1497	ZRLCC*GFOKSCRSRQCKOHRC*C-NH ₂	C5–C21
	2838.3	4	1554	ZRLC*C*GFOKSCRSRQCKOHRC*C*-NH ₂	C11–C16
1.6	2724.2	2	1440 + 1497	ZRLC*C'GFOKSCRSRQCKOHRC'C*-NH ₂	C4–C22 and C5–C21
	2838.3	4	1497	ZRLC*C*GFOKSCRSRQCKOHRC*C*-NH ₂	C11–C16
3.2	2610.2	0	1440	ZRLCCGFOKSCRSRQCKOHRCC-NH ₂	-
Disulfide connectivity with highest score: C4–C22, C5–C21, C11–C16					
[a] HPLC chromatogram, 10% eluent B for 15 min, then linear increase of eluent B up to 35% for another 25 min with A) 0.1% TFA in water and B) 0.1% TFA in acetonitrile. Retention time is given in minutes relative to the 6x-carbamidomethylated peptide version. [b] "Internal fragment found" is the fragment observed after MS/MS analysis for spectra containing the partial sequence SCRSRQCKOHRC with respective carbamidomethylated positions.					

Supporting Table 4. Internal fragments for 2- and 4-times carbamidomethylated (CAM) μ -PIIIA species with distinct disulfide bond pattern. The positions of the CAM modification indicate the disulfide linkage directly for 2-CAM, while in the case of 4-CAM the non-carbamidomethylated positions are connected.

Internal fragment (2-CAM) 2724	Resulting possible disulfide bond	Internal fragment (4-CAM) 2838	Resulting possible disulfide bond
1440	C4-C5	1497	C11,C16,C21,C22
	C4-C22		C5,C11,C16,C21
	C5-C22		C4,C11,C16,C21
1497	C4-C11	1554	C5,C16,C21,C22
	C4-C16		C5,C11,C21,C22
	C4-C21		C5,C11,C16,C22
	C5-C11		C4,C16,C21,C22
	C5-C16		C4,C11,C21,C22
	C5-C21		C4,C11,C16,C22
	C11-C22		C4,C5,C16,C21
	C16-C22		C4,C5,C11,C21
	C21-C22		C4,C5,C11,C16
1554	C11-C16	1611	C4,C5,C21,C22
	C11-C21		C4,C5,C16,C22
	C16-C21		C4,C5,C11,C22

Supporting Table 5. Overview of NMR calculations including alternative calculations for isomers possible from disulfide scrambling.

No.	Connectivity	1 st bond	2 nd bond	3 rd bond	Isomer found	Alternative isomers	NMR structure calculation (rmsd) for intended isomer	Structure calculation for alternative isomer	Structure calculation for alternative isomer
1*	C4-C21, C5-C22, C11-C16	C4-C21	C5-C22	C11- C16	1	11, 12	*1.46	12: 1.56 (0.75, 4-22)	11: 1.1 (0.65, 4-22)
2*	C4-C16, C5-C21, C11-C22	C4-C16	C11- C22	C5-C21	2	9, 10	*1.22	9: 1.2 (0.72, 4-22)	10: 0.94 (0.65, 4-22)
3*	C4-C5, C11-C16, C21-C22	C4-C5	C11- C16	C21- C22	3	4, 5	1.9	4: 3.5**	5: 3.9**
4	C4-C5, C11-C21, C16-C22	C4-C5	C11- C21	C16- C22	4	3, 5	2.4	3: 3.2 (0.7, 8-20)	5: 3.6 (1.0, 8-22)
5	C4-C5, C11-C22, C16-C21	C4-C5	C11- C22	C16- C21	5	3, 4	2.3	3: 4.8 (1.2, 7-18)	4: 3.4 (1.3 8-17)
6	C4-C11, C5-C16, C21-C22	C21- C22	C5-C16	C4-C11	6	3, 9	2.7 (1.2, 1-18)	3: 3.4 (1.9, 1-18)	9: 3.5 (1.9, 1-18)
7	C4-C11, C5-C21, C16-C22	C4-C11	C5-C21	C16- C22	-	6, 8	-	-	-
8	C4-C11, C5-C22, C16-C21	C16- C21	C5-C22	C4-C11	8	5, 13	1.5 (1.1, 4-22)	5: 4.4 (2.8, 4-22; 1.2, 10-22)	13: 2.6 (2.0, 4-22; 1.7, 4- 18)
9	C4-C16, C5-C11, C21-C22	C21- C22	C4-C16	C5-C11	9	3, 6	2.2 (1.3, 1-18)	3: 3.5	6: 3.1 (2.4, 1-18)
10	C4-C16, C5-C22, C11-C21	C11- C21	C5-C22	C4-C16	10	4, 14	1.5 (1.05, 4-22)	4: 2.2 (1.7, 4-22; 1.2, 4-18)	14: 2.2 (1.5, 4-22; 1.2, 4- 18)

11	C4-C21, C5-C11, C16-C22	C5-C11	C4-C21	C16- C22	11	9, 13	0.9	9: 3.4**	13: 4.1**
12	C4-C21, C5-C16, C11-C22	C5-C16	C11- C22	C4-C21	-	6, 14	-	-	-
13	C4-C22, C5-C11, C16-C21	C4-C22	C16- C21	C5-C11	-	14, 15	-	-	-
14	C4-C22, C5-C16, C11-C21	C4-C22	C11- C21	C5-C16	14	13, 15	0.9	13: 1.5	15: 2.0 (1.5, 4-22)
15	C4-C22, C5-C21, C11-C16	C4-C22	C5-C21	C11- C16	15	13, 14	2.2 (1.5, 4-22)	13: 4.7**	14: 4.7**

*These isomers have been described previously and were confirmed herein.⁷ **No convergence.

Supporting Table 6. Energy and structural statistics of the NMR ensembles of the μ -PIIIA isomers (lowest energy structures).

NMR distance & dihedral constraints (20 structures)												
Isomer	1*	2*	3	4	5	6	8	9	10	11	14	15
Distance constraints	137	218	204	204	212	192	190	161	209	185	214	237
Dihedral angle restraints	30	42	30	36	36	36	36	36	35	36	36	36
Structure statistics												
Violations	Mean (s.d.) in kJ/mol											
Total restraint violation energy			13.659	3.213	2.129	3.921	3.198	-9187	9.950	6.019	8.171	9.636
Distance restraints			13.603	3.138	1.961	3.710	3.123	2.584	9.894	5.959	5.293	9.532
Dihedral angle restraints	0.068	0.087	0.056	0.076	0.168	0.211	0.075	0.098	0.055	0.060	2.879	0.104
Total Force field energy (Yasara)	-233	-223	-9978	-9267	-9754	-8756	-9551	-9190	-10208	-9075	-7386	-10182
Internal solute energy			-2599	-2819	-2789	-1420	-2876	-1884	-2707	-2901	-1608	-2799
Electrostatic solv. energy			-7498	-6484	-7009	-7470	-6763	-7413	-7679	-6241	-5873	-7456
Van der Waals solv. energy			118	36	44	134	89	107	177	66	95	73

Mean global
r.m.s.d. (Å)

Backbone global (residues 1-22)	1.46 (±0.44)	1.22 (±0.42)	1.99 (±0.65)	2.46 (±0.76)	2.32 (±0.77)	2.77 (±1.24)	1.48 (±0.33)	2.12 (±0.54)	1.54 (±0.23)	0.94 (±0.40)	0.90 (±0.35)	2.21 (±0.78)
Heavy atoms (residues 1-22)	2.65 (±0.57)	2.31 (±0.46)	2.39 (±0.52)	2.85 (±0.70)	2.76 (±0.59)	3.09 (±1.17)	2.12 (±0.34)	2.49 (±0.55)	2.23 (±0.31)	1.49 (±0.31)	1.74 (±0.32)	2.70 (±0.68)
Backbone rigid (residues)	0.84 (4-22) (±0.24)	0.61 (4-22) (±0.22)	0.58 (6-19) (±0.16)	1.61 (4-22) (±0.34)	1.84 (4-22) (±0.64)	1.24 (1-18) (±0.50)	1.12 (4-22) (±0.26)	1.12 (1-18) (±0.41)	0.82 (3-17) (±0.21)	0.76 (4-22) (±0.31)	0.58 (4-22) (±0.13)	1.48 (4-22) (±0.33)
Heavy atoms rigid (residues)	1.88 (4-22) (±0.42)	1.61 (4-22) (±0.34)	1.16 (6-19) (±0.22)	2.02 (4-22) (±0.29)	2.03 (4-22) (±0.50)	1.86 (1-18) (±0.52)	1.80 (4-22) (±0.31)	1.67 (1-18) (±0.44)	1.57 (3-17) (±0.40)	1.45 (4-22) (±0.33)	1.27 (4-22) (±0.16)	1.94 (4-22) (±0.32)

*Data obtained from CYANA according to Tietze *et al.*⁷.

Supporting Table 7. NMR resonance assignments for structural isomers of μ -PIIIA for the respective isomers: proton chemical shifts (ppm) in H₂O/D₂O (9:1). The isomers 1 and 2 have been described previously in Tietze *et al.*⁷.

Isomer 3

	NH	H α	H β	H γ	H δ	H ϵ	H ζ
1 Glu	-	4.21	1.88, 2.38	2.27, 2.26	-	-	-
2 Arg	8.31	4.17	1.68, 1.62	1.52, 1.51	3.06	-	-
3 Leu	8.33	4.25	1.49	1.49	0.70, 0.76	-	-
4 Cys	8.61	4.15	3.26, 3.56	-	-	-	-
5 Cys	7.93	4.88	2.59, 3.00	-	-	-	-
6 Gly	8.5	3.75, 3.58	-	-	-	-	-
7 Phe	8.09	4.74	2.99, 2.78	-	7.13	-	7.2
8 Hyp	-	4.44	1.92, 2.20	4.46	3.57, 3.76	-	-
9 Lys	8.53	4.16	1.67, 1.74	1.38, 1.36	1.58	2.88	-
10 Ser	8.14	4.27	3.74, 3.72	-	-	-	-
11 Cys	8.32	4.59	2.99, 3.07	-	-	-	-
12 Arg	8.63	4.31	1.60, 1.85	1.49	3.07	7.05	-
13 Ser	7.78	4.37	3.89, 3.78	-	-	-	-
14 Arg	8.3	4.02	1.75	1.52	3.08	7.07	-
15 Gln	8.1	4.17	1.85, 1.96	2.21	-	-	-
16 Cys	7.97	4.46	2.96, 3.08	-	-	-	-
17 Lys	8.06	4.45	1.58, 1.66	1.30, 1.34	1.54	2.84	7.42
18 Hyp	-	4.39	1.83, 2.17	4.45	3.64, 3.73	-	-
19 His	8.67	4.53	3.09	-	-	-	-
20 Arg	8.36	4.18	1.60, 1.65	1.46	3.04	7.06	-
21 Cys	8.16	4.81	3.48, 3.28	-	-	-	-
22 Cys	8.67	3.82	3.19	-	-	-	-

Isomer 4

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.77	4.21	2.38	-	-	-	-
2	Arg	8.26	4.19	1.60, 1.66	1.51	3.07	-	-
3	Leu	8.33	4.25	1.48, 1.64	-	0.72	-	-
4	Cys	8.61	4.15	3.55, 3.27	-	-	-	-
5	Cys	7.94	4.91	3.31, 2.59	-	-	-	-
6	Gly	8.5	3.76, 3.75	-	-	-	-	-
7	Phe	8.1	4.75	2.99, 2.77	-	-	-	-
8	Hyp	-	4.45	1.92, 2.20	4.46	3.76, 3.57	-	-
9	Lys	8.53	4.14	1.74, 1.67	1.36, 1.39	1.58	2.89	-
10	Ser	8.15	4.28	3.74, 3.71	-	-	-	-
11	Cys	8.31	4.6	3.06, 2.99	-	-	-	-
12	Arg	8.62	4.31	1.60, 1.84	1.48	3.06	7.06	-
13	Ser	7.79	4.36	3.78, 3.89	-	-	-	-
14	Arg	8.29	4.02	-	-	3.07	-	-
15	Gln	8.1	4.17	1.96, 1.85	2.21	-	7.06	-
16	Cys	7.97	4.46	2.94, 3.08	-	-	-	-
17	Lys	8.08	4.45	1.58, 1.66	1.33, 1.32	1.54	2.84	-
18	Hyp	-	4.39	2.18, 1.83	4.45	3.72, 3.64	-	-
19	His	8.63	4.52	3.07	-	-	-	-
20	Arg	8.33	4.18	1.66, 1.75	1.49	3.07	-	-
21	Cys	8.16	-	3.28, 3.48	-	-	-	-
22	Cys	8.68	3.81	-	-	-	-	-

Isomer 5

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.76	4.22	2.39, 1.88	2.27	-	-	-
2	Arg	8.33	4.2	1.74, 1.65	1.51	3.06	-	-
3	Leu	8.33	4.24	1.50	-	0.70, 0.76	-	-
4	Cys	8.61	4.16	3.27	-	-	-	-
5	Cys	7.91	4.89	3.30, 2.60	-	-	-	-
6	Gly	8.5	3.75	-	-	-	-	-
7	Phe	8.09	4.74	2.99, 2.77	-	7.13	-	-
8	Hyp	-	4.45	1.92, 2.20	4.45	3.58, 3.76	-	-
9	Lys	8.53	4.15	1.74, 1.66	1.39, 1.38	1.58	2.88	-
10	Ser	8.15	4.27	3.72, 3.73	-	-	-	-
11	Cys	8.31	4.59	2.99, 3.07	-	-	-	-
12	Arg	8.62	4.31	1.59, 1.85	1.50, 1.49	3.06, 2.98	-	-
13	Ser	7.78	4.37	3.89, 3.78	-	-	-	-
14	Arg	8.3	4.02	1.75	1.52	3.07	-	-
15	Gln	8.1	4.17	1.96, 1.85	2.21	-	-	-
16	Cys	7.97	4.46	3.08, 2.96	-	-	-	-
17	Lys	8.06	4.45	1.66, 1.58	1.33, 1.32	1.53	2.84	-
18	Hyp	-	4.39	1.83, 2.17	4.45	3.64, 3.72	-	-
19	His	8.67	4.53	3.09	-	-	-	-
20	Arg	8.36	4.2	1.66, 1.60	1.46	3.04	-	-
21	Cys	8.15	4.81	3.28, 3.48	-	-	-	-
22	Cys	8.67	3.81	3.21	-	-	-	-

Isomer 6

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.77	4.23	-	-	-	-	-
2	Arg	8.32	4.19	1.59, 1.66	1.50, 1.51	-	-	-
3	Leu	8.33	4.27	1.53	1.45	0.74	-	-
4	Cys	8.61	4.16	3.57, 3.28	-	-	-	-
5	Cys	7.93	4.89	2.59	-	-	-	-
6	Gly	8.5	3.76	-	-	-	-	-
7	Phe	8.09	4.74	2.78, 3.00	-	7.14	-	-
8	Hyp	-	4.46	2.21, 1.92	4.46	3.59, 3.76	-	-
9	Lys	8.53	4.16	1.67, 1.75	1.40	1.58	-	7.42
10	Ser	8.15	4.28	3.72, 3.75	-	-	-	-
11	Cys	8.32	4.59	3.08, 2.99	-	-	-	-
12	Arg	8.63	4.31	1.85, 1.60	1.51	3.07	-	-
13	Ser	7.79	4.37	3.90, 3.79	-	-	-	-
14	Arg	8.3	4.03	1.76	1.53	3.08	-	-
15	Gln	8.1	4.18	1.95, 1.86	2.22	-	-	-
16	Cys	7.98	4.47	2.96, 3.08	-	-	-	-
17	Lys	8.07	4.46	1.59, 1.67	1.33	1.55	-	-
18	Hyp	-	4.4	2.18, 1.84	4.46	3.65, 3.73	-	-
19	His	8.65	4.53	3.08	-	-	-	-
20	Arg	8.34	4.19	1.71, 1.61	1.5	-	-	-
21	Cys	8.16	4.82	3.48, 3.29	-	-	-	-
22	Cys	8.67	3.82	3.22	-	-	-	-

Isomer 8

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.77	4.23	2.27, 2.38	1.87	-	-	-
2	Arg	8.33	4.17	1.64, 1.74	1.47	-	-	-
3	Leu	8.34	4.25	1.40, 1.53	-	0.72	-	-
4	Cys	8.61	4.15	3.26, 3.55	-	-	-	-
5	Cys	7.9	4.88	3.28, 2.59	-	-	-	-
6	Gly	8.5	3.75	-	-	-	-	-
7	Phe	8.09	4.74	2.77, 2.99	-	7.13	-	-
8	Hyp	-	4.45	1.93, 2.20	4.46	3.58, 3.76	-	-
9	Lys	8.53	4.15	1.66, 1.74	1.36	1.57	2.88	-
10	Ser	8.15	4.26	3.71, 3.73	-	-	-	-
11	Cys	8.31	4.59	2.98, 3.07	-	-	-	-
12	Arg	8.62	4.31	1.84, 1.59	1.47	3.07	-	-
13	Ser	7.78	4.37	3.78, 3.89	-	-	-	-
14	Arg	8.3	4.02	1.75	1.52	3.07	-	-
15	Gln	8.1	4.17	1.96, 1.85	2.21	-	-	-
16	Cys	7.97	4.46	2.95, 3.08	-	-	-	-
17	Lys	8.06	4.45	1.66, 1.58	1.33, 1.29	1.53	2.84	-
18	Hyp	-	4.4	2.16, 1.83	4.45	3.72, 3.64	-	-
19	His	8.67	4.53	3.09	-	-	-	-
20	Arg	8.36	4.18	1.66, 1.59	1.45	3.08	-	-
21	Cys	8.16	4.82	3.47, 3.28	-	-	-	-
22	Cys	8.67	3.81	3.2	-	-	-	-

Isomer 9

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	-	4.24	2.38, 1.88	2.28	-	-	-
2	Arg	8.33	4.17	1.68, 1.62	1.47, 1.45	3.05	-	-
3	Leu	8.32	4.26	1.49	-	0.73	-	-
4	Cys	8.61	4.14	3.54, 3.26	-	-	-	-
5	Cys	7.93	4.88	2.61	-	-	-	-
6	Gly	8.5	3.75	-	-	-	-	-
7	Phe	8.09	4.74	2.99, 2.77	-	7.13	-	-
8	Hyp	-	4.44	1.92, 2.20	4.45	3.58, 3.75	-	-
9	Lys	8.53	4.15	1.74, 1.66	1.34, 1.39	1.58	2.88	-
10	Ser	8.15	4.27	3.74, 3.71	-	-	-	-
11	Cys	8.32	4.59	3.07, 2.98	-	-	-	-
12	Arg	8.62	4.31	1.84, 1.60	1.49	3.06	-	-
13	Ser	7.78	4.36	3.89, 3.78	-	-	-	-
14	Arg	8.3	4.02	1.75	1.53	3.07	-	-
15	Gln	8.1	4.17	1.85, 1.96	2.21	-	-	-
16	Cys	7.97	4.46	2.95, 3.08	-	-	-	-
17	Lys	8.07	4.45	1.66, 1.58	1.34, 1.3	1.54	2.84	-
18	Hyp	-	4.39	2.17, 1.83	4.45	3.73, 3.64	-	-
19	His	8.66	4.52	3.07	-	7.16	-	-
20	Arg	8.35	4.18	1.66, 1.59	1.47	3.05	-	-
21	Cys	8.16	4.81	3.48	-	-	-	-
22	Cys	8.67	3.81	-	-	-	-	-

Isomer 10

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.77	4.23	2.38, 1.86	2.27	-	-	-
2	Arg	8.33	4.18	1.74, 1.64	1.51	-	-	-
3	Leu	8.32	4.25	1.46, 1.47	-	0.77, 0.70	-	-
4	Cys	8.61	4.15	3.54, 3.26	-	-	-	-
5	Cys	7.93	4.88	3.00, 2.91	-	-	-	-
6	Gly	8.5	3.75	-	-	-	-	-
7	Phe	8.09	4.73	2.77, 2.99	-	7.18	-	-
8	Hyp	-	4.45	1.92, 2.20	4.46	3.57, 3.76	-	-
9	Lys	8.53	4.15	1.74, 1.66	1.35, 1.40	1.57	2.87	-
10	Ser	8.15	4.27	3.73	-	-	-	-
11	Cys	8.32	4.59	3.07, 2.99	-	-	-	-
12	Arg	8.63	4.31	1.85, 1.59	1.49	3.06	-	-
13	Ser	7.78	4.37	3.89, 3.78	-	-	-	-
14	Arg	8.3	4.02	1.75, 1.76	1.51, 1.53	3.07	-	-
15	Gln	8.1	4.17	1.85, 1.96	2.21	-	-	-
16	Cys	7.97	4.46	2.96, 3.08	-	-	-	-
17	Lys	8.07	4.45	1.66, 1.58	1.33, 1.3	1.52	2.84	-
18	Hyp	-	4.4	2.16, 1.83	4.45	3.72, 3.64	-	-
19	His	8.67	4.53	3.08	-	-	-	-
20	Arg	8.36	4.18	1.66, 1.59	1.47, 1.48	3.07	-	-
21	Cys	8.16	4.82	3.48, 3.28	-	-	-	-
22	Cys	8.67	3.81	-	-	-	-	-

Isomer 11

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.80	4.3	2.46	2.35	-	-	-
2	Arg	8.35	4.22	1.74, 1.70	1.53	3.12	-	-
3	Leu	8.3	4.35	1.60, 1.53	-	0.78, 0.84	-	-
4	Cys	8.44	4.6	3.02, 3.13	-	-	-	-
5	Cys	8.28	4.65	3.05, 3.17	-	-	-	-
6	Gly	8.39	3.79, 3.71	-	-	-	-	-
7	Phe	-	4.83	2.93, 3.17	-	-	-	7.27
8	Hyp	-	4.41	2.03, 2.28	4.57	3.83, 3.80	-	-
9	Lys	8.42	4	1.81	1.37	1.61	2.91	-
10	Ser	7.89	4.42	3.87, 3.80	-	-	-	-
11	Cys	8.13	4.66	3.19, 3.10	-	-	-	-
12	Arg	8.31	4.29	1.75, 1.74	-	1.53-	-	-
13	Ser	7.98	4.42	3.79, 3.84	-	-	-	-
14	Arg	8.21	4.31	1.71, 1.84	1.58	3.13	-	-
15	Gln	8.29	4.46	1.91, 2.04	2, 27	-	-	-
16	Cys	8.69	4.59	-	-	-	-	-
17	Lys	8.38	4.46	1.77, 1.75	1.4	1.50	-	-
18	Hyp	-	4.29	1.99, 2.27	4.53	3.69, 3.68	-	-
19	His	8.39	4.64	3.05, 3.17	-	7.14	-	8.55-
20	Arg	8.44	4.0	1.8	1.6	3.13	-	-
21	Cys	8.51	4.78	3.08, 3.01	-	-	-	-
22	Cys	8.74	4.75	3.21, 3.06	-	-	-	-

Isomer 13

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	-	4.21	1.88, 2.38	2.27, 2.26	-	-	-
2	Arg	8.31	4.17	1.68, 1.62	1.52, 1.51	3.06	-	-
3	Leu	8.33	4.25	1.49	1.49	0.70, 0.76	-	-
4	Cys	8.61	4.15	3.26, 3.56	-	-	-	-
5	Cys	7.93	4.88	2.59, 3.00	-	-	-	-
6	Gly	8.5	3.75, 3.58	-	-	-	-	-
7	Phe	8.09	4.74	2.99, 2.78	-	7.13	-	7.2
8	Hyp	-	4.44	1.92, 2.20	4.46	3.57, 3.76	-	-
9	Lys	8.53	4.16	1.67, 1.74	1.38, 1.36	1.58	2.88	-
10	Ser	8.14	4.27	3.74, 3.72	-	-	-	-
11	Cys	8.32	4.59	2.99, 3.07	-	-	-	-
12	Arg	8.63	4.31	1.60, 1.85	1.49	3.07	7.05	-
13	Ser	7.78	4.37	3.89, 3.78	-	-	-	-
14	Arg	8.3	4.02	1.75	1.52	3.08	7.07	-
15	Gln	8.1	4.17	1.85, 1.96	2.21	-	-	-
16	Cys	7.97	4.46	2.96, 3.08	-	-	-	-
17	Lys	8.06	4.45	1.58, 1.66	1.30, 1.34	1.54	2.84	7.42
18	Hyp	-	4.39	1.83, 2.17	4.45	3.64, 3.73	-	-
19	His	8.67	4.53	3.09, 3.09	-	-	-	-
20	Arg	8.36	4.18	1.60, 1.65	1.46	3.04	7.06	-
21	Cys	8.16	4.81	3.48, 3.28	-	-	-	-
22	Cys	8.67	3.82	3.19	-	-	-	-

Isomer 14

		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.77	4.22	1.89, 2.40	2.27	-	-	-
2	Arg	8.31	4.17	1.68, 1.62	1.51, 1.50	3.06	-	-
3	Leu	8.33	4.25	1.48	1.49	0.71, 0.77	-	-
4	Cys	8.61	4.15	3.54, 3.26	-	-	-	-
5	Cys	7.98	4.9	3.39, 2.68	-	-	-	-
6	Gly	8.5	3.76	-	-	-	-	-
7	Phe	8.09	4.74	2.99, 2.78	-	7.13	-	-
8	Hyp	-	4.44	2.20, 1.92	4.46	3.76, 3.58-	-	-
9	Lys	8.53	4.16	1.67, 1.74	1.35, 1.39	1.58	2.89	-
10	Ser	8.13	4.27	3.74, 3.72	-	-	-	-
11	Cys	8.32	4.59	2.99, 3.07	-	-	-	-
12	Arg	8.62	4.31	1.60, 1.85	1.49	3.06, 3.07	-	-
13	Ser	7.78	4.37	3.89, 3.78	-	-	-	-
14	Arg	8.3	4.02	1.75	1.52	3.08	7.07	-
15	Gln	8.1	4.17	1.96, 1.86	2.21	-	-	-
16	Cys	7.97	4.46	2.96, 3.08	-	-	-	-
17	Lys	8.06	4.45	1.58, 1.66	1.34, 1.30	1.53, 1.54	2.84	-
18	Hyp	-	4.39	2.17, 1.83	4.45	3.64, 3.72	-	-
19	His	8.67	4.53	3.09	-	-	-	-
20	Arg	8.36	4.18	1.60, 1.65	1.46	3.04	-	-
21	Cys	8.16	4.82	3.28, 3.48	-	-	-	-
22	Cys	8.67	3.81	3.2	-	-	-	-

Isomer 15

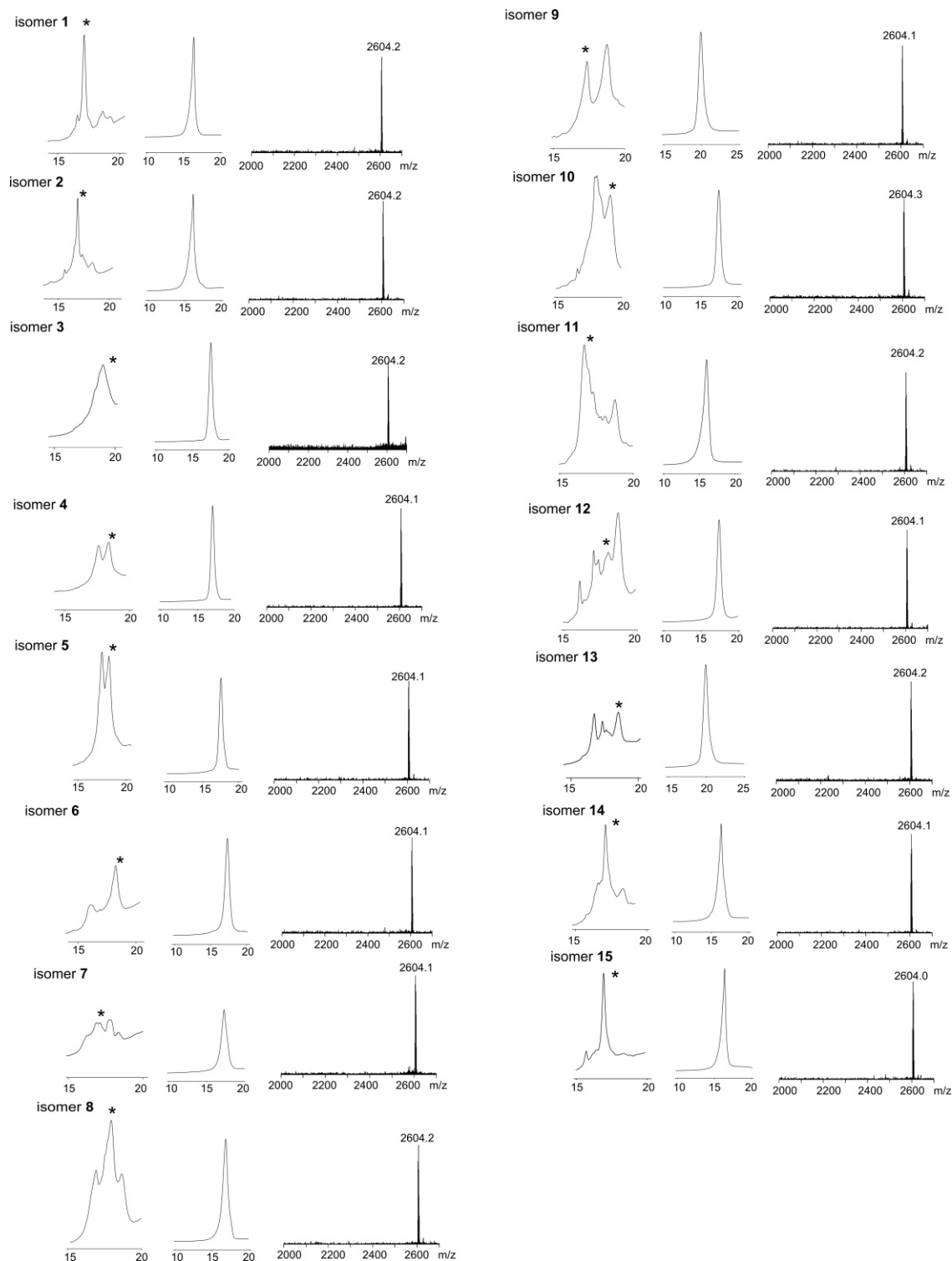
		NH	H α	H β	H γ	H δ	H ϵ	H ζ
1	Glu	7.78	4.24	-	-	-	-	-
2	Arg	8.32	4.19	1.71, 1.64	1.52, 1.48	3.06	7.08	-
3	Leu	8.38	4.28	1.45, 1.48	1.55	0.72, 0.78	-	-
4	Cys	8.42	4.53	3.14	-	-	-	-
5	Cys	8.29	4.59	3.12, 3.17	-	-	-	-
6	Gly	8.13	3.68, 3.78	-	-	-	-	-
7	Phe	7.99	4.73	2.98, 2.79	-	7.13	-	-
8	Hyp	-	4.47	2.22, 1.93	4.47	3.73, 3.60	-	-
9	Lys	8.49	4.13	1.68, 1.74	1.35, 1.38	1.57	2.87	7.42
10	Ser	8.09	4.26	3.72, 3.76	-	-	-	-
11	Cys	8.23	4.58	3.08, 2.97	-	-	-	-
12	Arg	8.63	4.35	1.87, 1.59	1.49, 1.52	3.07, 3.08	7.06	-
13	Ser	7.76	4.41	3.78, 3.91	-	-	-	-
14	Arg	8.33	4.05	1.76, 1.78	1.54	3.08	7.1	-
15	Gln	8.1	4.22	1.84, 1.97	2.21	-	-	-
16	Cys	7.94	4.44	2.92, 3.04	-	-	-	-
17	Lys	8.2	4.42	1.64	1.26, 1.32	1.57	2.82	-
18	Hyp	-	4.41	1.84, 2.17	4.47	3.73, 3.63	-	-
19	His	8.53	4.55	3.06, 3.10	-	-	-	-
20	Arg	8.44	4.2	1.74, 1.65	1.51	3.09, 3.06	7.06	-
21	Cys	8.52	4.66	3.10, 3.15	-	-	-	-
22	Cys	8.22	4.53	3.25, 3.14	-	-	-	-

Supporting Table 8. Summary of the HADDOCK docking runs. For each isomer docking the three best-scoring clusters are outlined. Docking results are ranked according to their Z-scores.

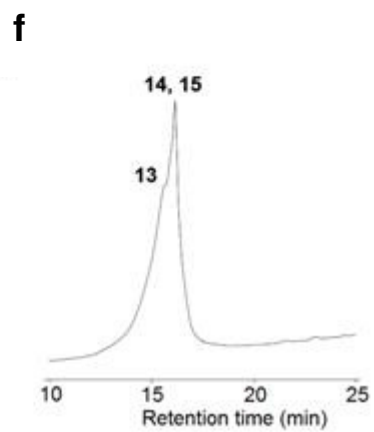
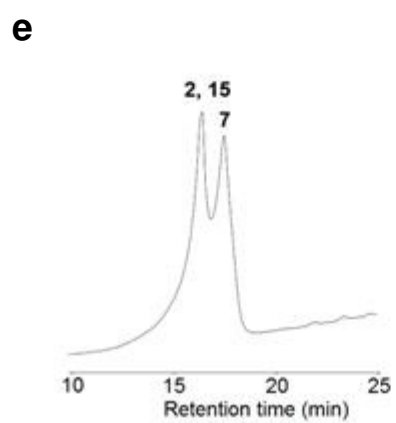
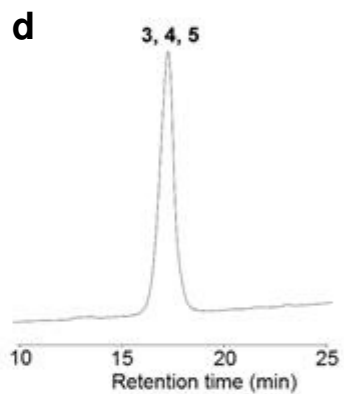
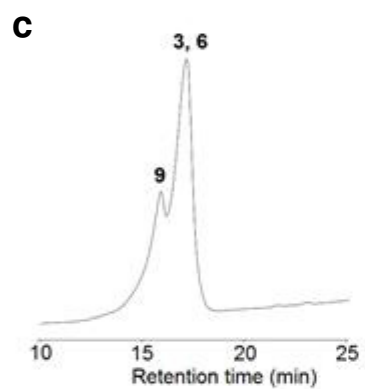
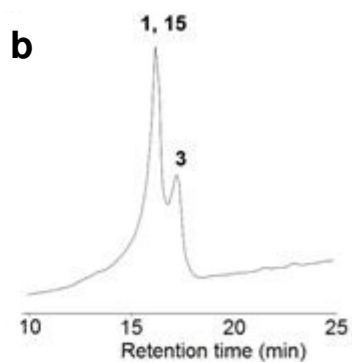
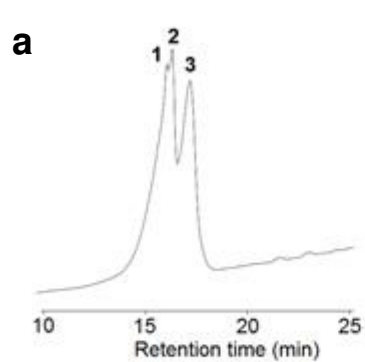
Isomer	Z-score	HADDOCK score	Cluster size	RMSD from the overall lowest-energy structure [Å]	Buried surface area [Å ²]	Vina score binding energy [kcal/mol]
1	-1.2	174.1 ± 6	25	3.2 ± 0.0	3114.0 ± 168	12.2
	-1.2	175.3 ± 17	19	0.7 ± 0.5	2888.9 ± 212	
	-0.9	181.9 ± 4	60	2.8 ± 0.1	2909.7 ± 72	
2	-1.4	162.9 ± 20	41	1.3 ± 0.8	2877.3 ± 134	10.4
	-0.9	182.3 ± 5	18	2.5 ± 0.1	2961.9 ± 146	
	-0.8	183.6 ± 13	26	2.2 ± 0.2	2817.6 ± 121	
3	-1.6	127.2 ± 11	45	1.2 ± 0.7	3349.9 ± 140	7.9
	-1.5	129.7 ± 2	92	3.7 ± 0.1	3435.1 ± 51	
	-0.8	151.2 ± 24	9	3.1 ± 0.0	3185.7 ± 201	
4	-2.1	139.5 ± 5	89	1.4 ± 0.9	3239.6 ± 43	10.7
	-0.9	181.0 ± 3	20	3.5 ± 0.3	3078.0 ± 98	
	-0.3	202.0 ± 5	33	3.3 ± 0.1	2906.4 ± 142	
5	-0.9	158.8 ± 3	41	3.2 ± 0.0	3237.1 ± 51	11.1
	-0.6	168.9 ± 11	55	3.1 ± 0.2	3038.7 ± 97	
	-0.6	169.6 ± 35	54	0.9 ± 0.6	3040.3 ± 44	
6	-1.8	132.2 ± 8	42	0.5 ± 0.3	3360.5 ± 98	11.5
	-1.0	156.1 ± 16	19	3.0 ± 0.1	3120.9 ± 124	
	-0.8	163.2 ± 11	19	2.9 ± 0.0	3254.4 ± 35	
7	-1.2	171.8 ± 7	38	0.7 ± 0.5	2913.6 ± 123	10.8
	-1.0	177.3 ± 10	40	2.5 ± 0.0	3028.1 ± 123	
	-0.6	187.5 ± 7	18	2.5 ± 0.1	3056.5 ± 164	
8	-1.4	151.5 ± 8	14	2.6 ± 0.1	3083.6 ± 141	12.2
	-1.0	162.9 ± 16	14	0.8 ± 0.5	3004.7 ± 138	
	-0.4	180.4 ± 8	58	2.4 ± 0.0	3120.0 ± 59	
9	-1.3	162.1 ± 20	19	1.2 ± 0.8	3262.9 ± 243	11.0
	-0.9	169.3 ± 7	28	2.3 ± 0.2	3168.8 ± 80	
	-0.9	170.1 ± 23	40	3.0 ± 0.1	2912.1 ± 270	
10	-1.6	142.2 ± 28	25	0.9 ± 0.7	3134.1 ± 52	11.4
	-0.9	165.1 ± 4	10	2.3 ± 0.1	3022.5 ± 110	
	-0.8	169.6 ± 18	28	2.2 ± 0.1	3057.6 ± 79	
11	-2.2	131.2 ± 21	28	0.5 ± 0.3	3019.3 ± 99	12.1
	-0.8	164.1 ± 14	24	3.0 ± 0.0	3123.3 ± 62	
	-0.4	174.4 ± 14	30	2.2 ± 0.1	2967.6 ± 34	
12	-1.9	123.5 ± 11	82	0.8 ± 0.5	3220.7 ± 55	11.1
	-0.7	169.6 ± 11	37	2.4 ± 0.0	3021.8 ± 110	
	-0.7	171.4 ± 14	21	2.8 ± 0.1	2903.0 ± 155	
13	-1.0	193.0 ± 19	22	2.8 ± 0.0	2913.5 ± 108	10.2
	-0.7	201.4 ± 29	78	1.1 ± 0.9	2858.1 ± 161	
	-0.6	204.5 ± 4	21	2.1 ± 0.0	2891.9 ± 44	
14	-1.8	181.0 ± 12	89	0.7 ± 0.5	2830.2 ± 41	10.8
	0	211.4 ± 6	45	2.5 ± 0.0	2713.1 ± 98	
	0	211.7 ± 21	34	2.0 ± 0.2	2679.1 ± 126	
15	-1.5	162.2 ± 21	65	1.2 ± 0.7	2966.4 ± 186	9.9
	-1.3	168.1 ± 12	28	2.0 ± 0.1	3096.8 ± 29	
	-0.9	179.5 ± 13	9	3.4 ± 0.1	3146.3 ± 147	

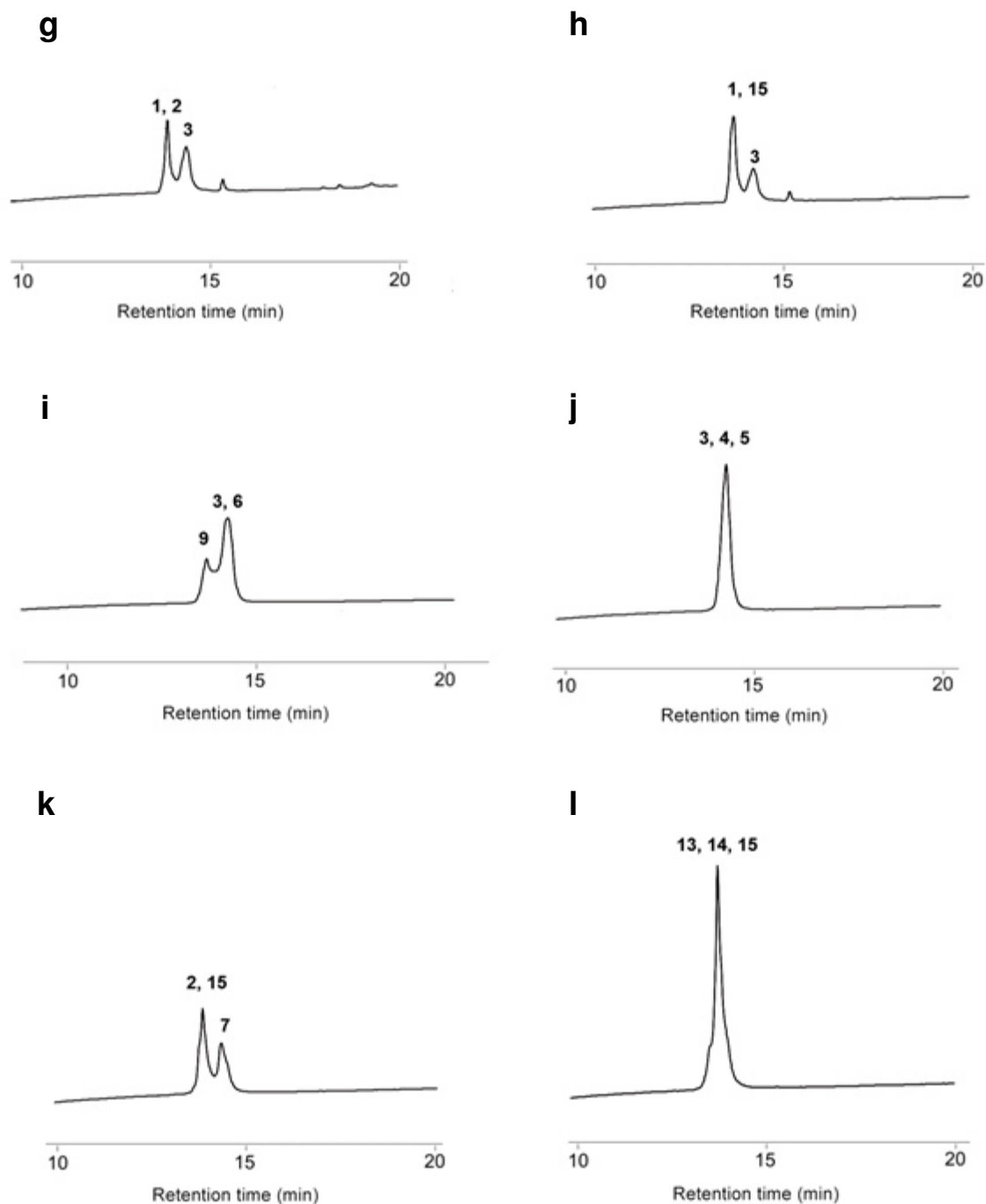
Supporting Table 9. H-bond analysis of the best scoring μ -PIIIA-channel complex

Isomer	total number of H-bonds	domain			
		I loops/SS1- SS2	II loops/SS1- SS2	III loops/SS1- SS2	IV loops/SS1- SS2
1	19	3/4	0/4	3/1	2/2
2	12	1/1	2/1	0/1	4/2
3	17	2/4	0/4	2/1	2/2
4	23	6/3	1/1	4/1	6/2
5	17	4/2	2/3	0/3	1/2
6	21	0/4	1/7	2/3	2/2
7	17	2/3	1/4	0/3	1/1
8	26	3/7	1/4	0/2	3/6
9	23	1/5	4/3	5/2	1/2
10	21	3/4	1/2	2/1	5/4
11	19	0/3	4/4	2/4	0/2
12	22	1/3	4/5	2/3	1/3
13	20	2/7	1/2	2/3	0/3
14	17	2/3	1/1	0/3	3/4
15	18	4/1	0/3	1/1	5/3

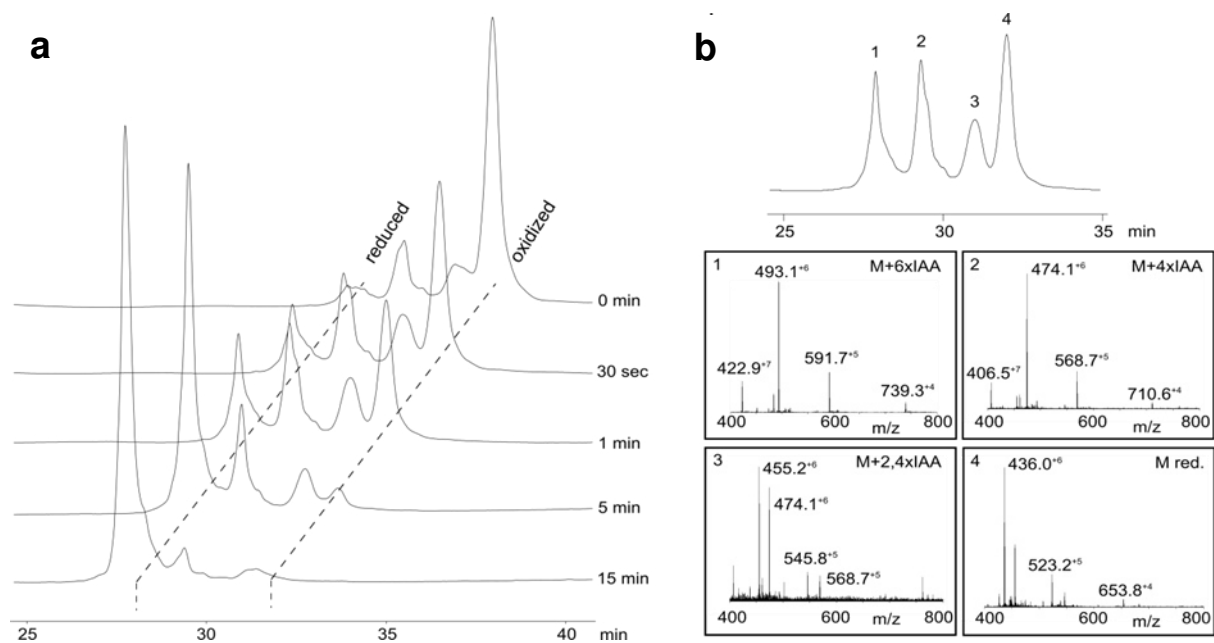


Supporting Figure 1. HPLC profiles of crude (left), and purified (center) μ -PIIIA isomers as well as corresponding MALDI-TOF mass spectra (right). Peak marked with an asterisks (*) is the fraction used for NMR structure analysis.

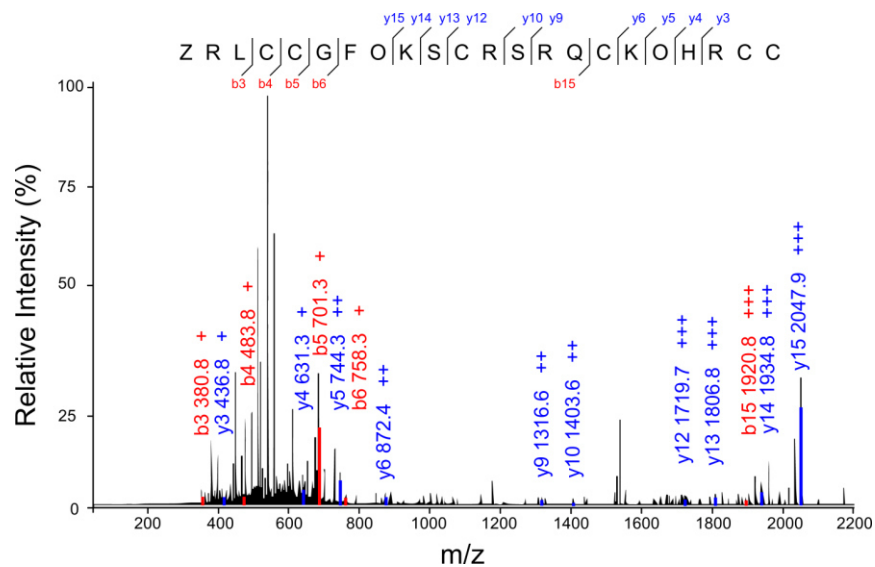




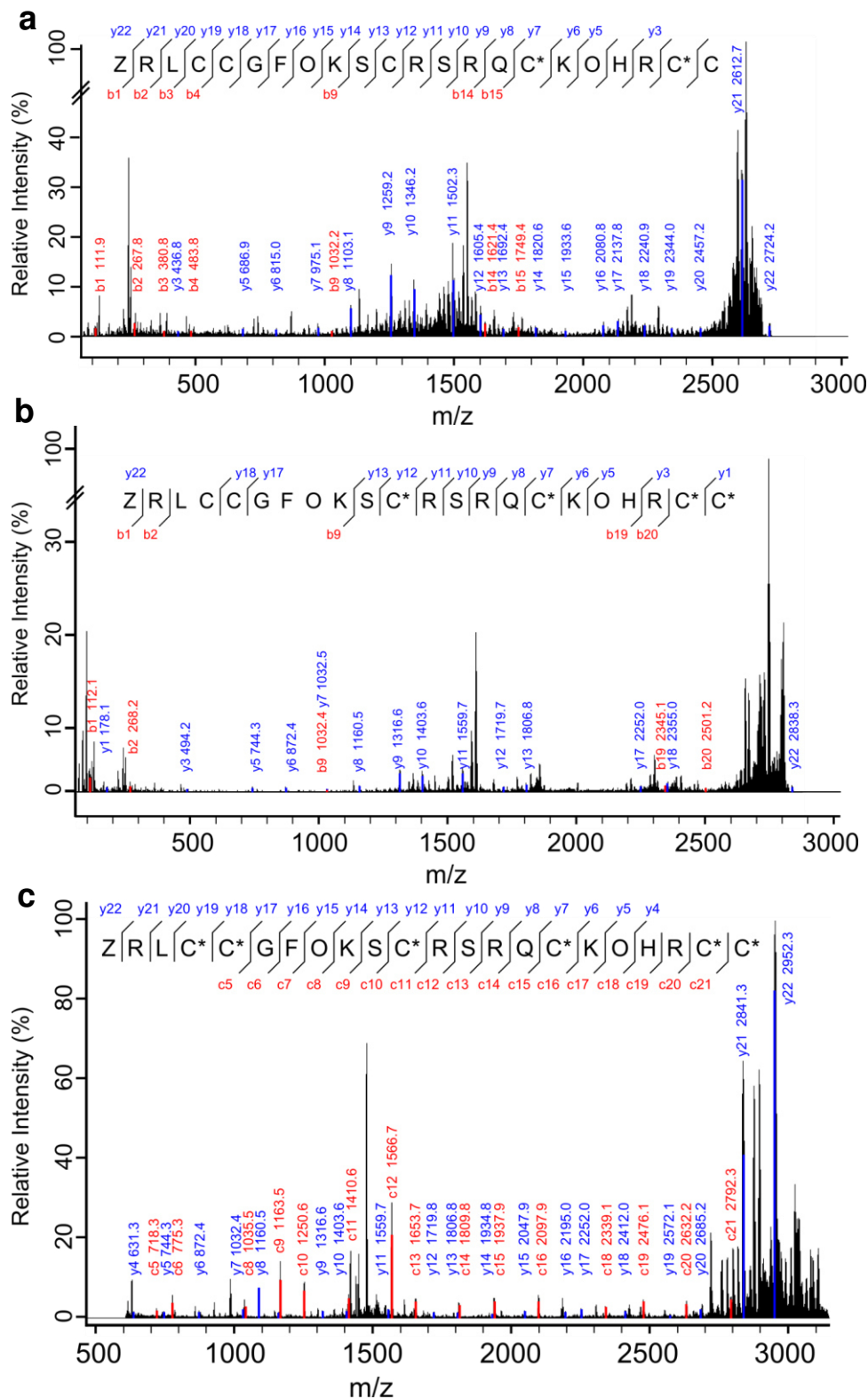
Supporting Figure 2. Co-elution profiles of different mixtures of μ -PiIIA isomers on a C18 column: a) isomers **1**, **2**, and **3** (originally isolated from an oxidative self-folding approach described earlier⁷), b) isomers **1**, **3**, and **15** sharing disulfide bridge C11–C16, c) isomers **3**, **6**, and **9** with common disulfide bond C21–C22, d) isomers **3**, **4**, and **5** with disulfide bond C4–C5, e) isomers **2**, **7**, and **15** sharing C5–C21, and f) isomers **13**, **14**, and **15** with common disulfide bond C4–C22. g-l) Co-elution profiles for the same mixtures g-l) separated on a C4 column.



Supporting Figure 3. a) HPLC elution profiles for partial reduction of μ -PIIIA isomer **3** after addition of TCEP and vigorously shaking for 10 sec. b) HPLC of reaction mixture 1 minute after addition of TCEP with respective masses found in each fraction (ESI-MS analysis).

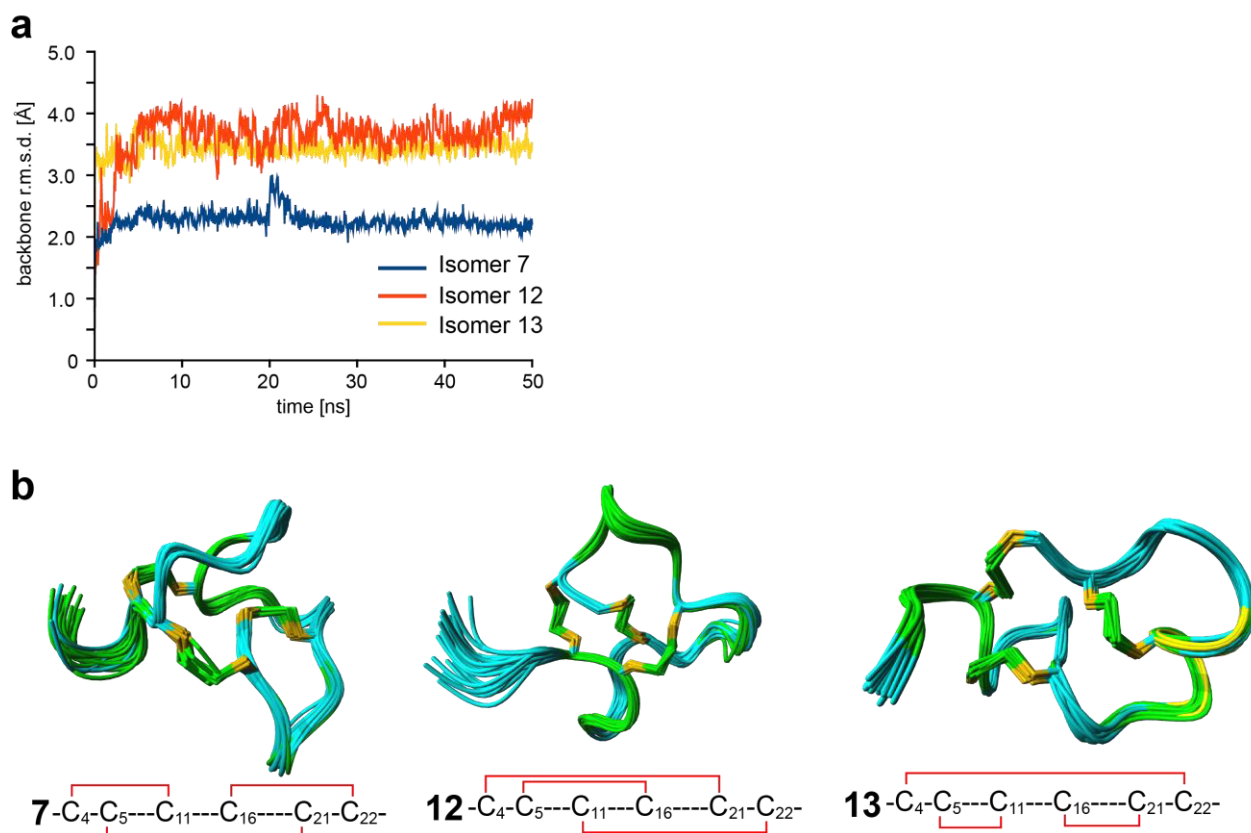


Supporting Figure 4. Sequence analysis of linear μ -PIIIA by LC-ESI-MS/MS indicating that the full-length peptide could not be analyzed due to insufficient fragmentation and low sequence coverage. “+” indicates the charge of the observed fragment ion.

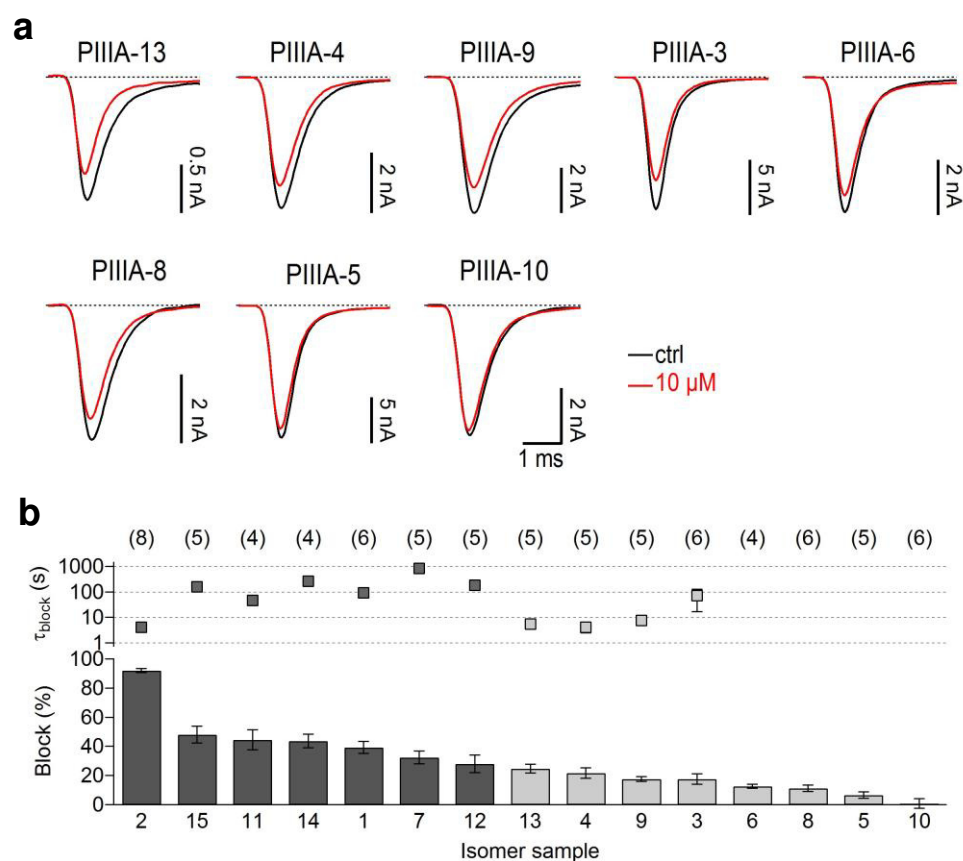


Supporting Figure 5. a) MALDI-MS/MS sequencing of a 2-times carbamidomethylated species of isomer **5** after fragmentation obtained with LID TOF/TOF. Cysteine residues 16 and 21 were carbamidomethylated. b) MALDI MS/MS LID TOF/TOF sequencing of the corresponding 4-times carbamidomethylated species. Cysteine residues 11, 16, 21 and 22 were carbamidomethylated

indicating the disulfide pattern of isomer **5** (C4-C5, C11-C22, C16-C21). c) MALDI ISD-MS/MS spectra obtained from 6-times carbamidomethylated peptide. Although the sequence coverage is higher compared to fragmentation with MS/MS LID TOF/TOF, no preselection of the precursor ion is possible which is essential for discrimination between different carbamidomethylated species present in one fraction. In this case c- and y-ions are formed in contrast to b- and y-ions by MS/MS LID TOF/TOF.

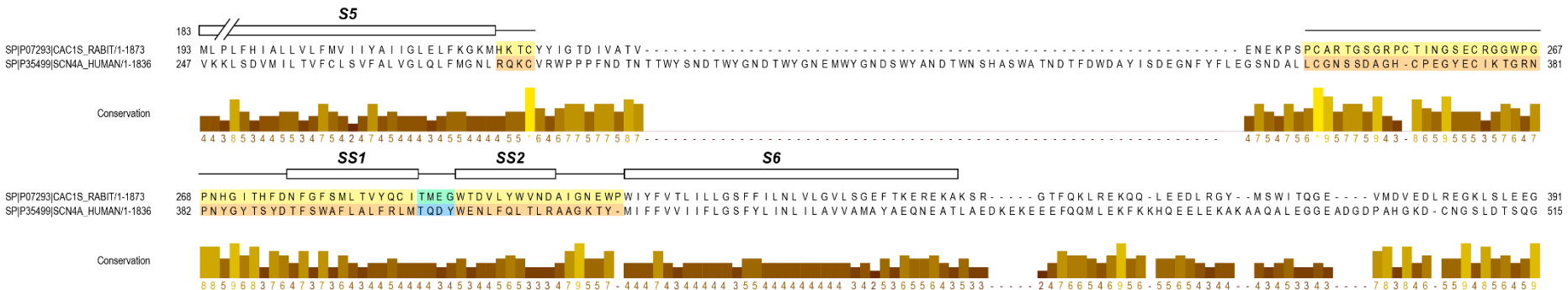


Supporting Figure 6. Structure predictions for μ -PIIIA isomers **7**, **12** and **13**. a) Development of the backbone r.m.s.d. during molecular dynamic simulations for the three isomers, indicating structural equilibration after ~ 30 ns of simulation. b) structural ensembles (20 structures) generated from the last 20 ns of the simulation (mean r.m.s.d, backbone global: **7** 0.61 Å, **12** 0.72 Å, **13** 0.41 Å)

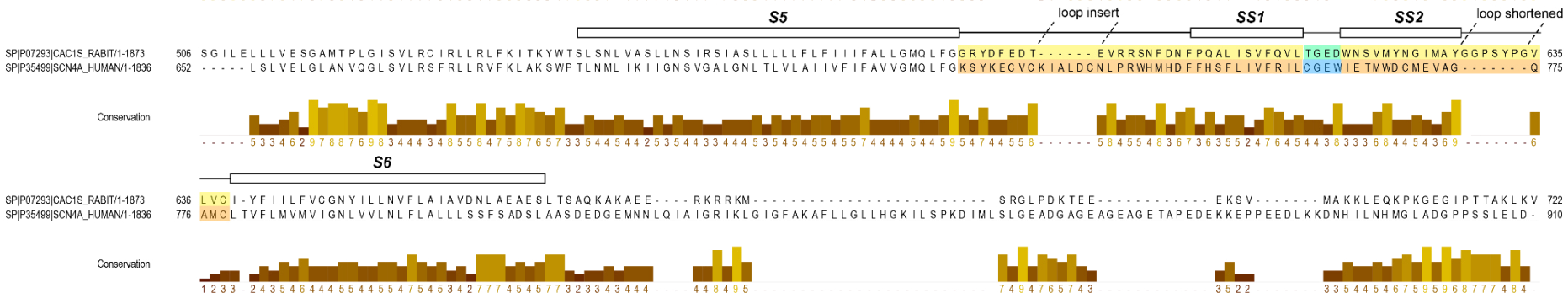


Supporting Figure 7. Block of $\text{Na}_v1.4$ -mediated currents by disulfide isomers of μ -PIIIA. a) Representative current traces of transiently expressed $\text{Na}_v1.4$ channels evoked at a test potential of -20 mV before (black, ctrl) and after (red) application of $10 \mu\text{M}$ of the indicated μ -PIIIA isomers. b) Steady-state block of $\text{Na}_v1.4$ -mediated currents by $10 \mu\text{M}$ of the indicated μ -PIIIA isomers (*bottom*) and the associated single exponential time constant, τ_{block} , describing the onset of block (*top*). Dark gray indicates isomers already shown in Figure 3. Numbers of individual experiments, n , are provided in parentheses.

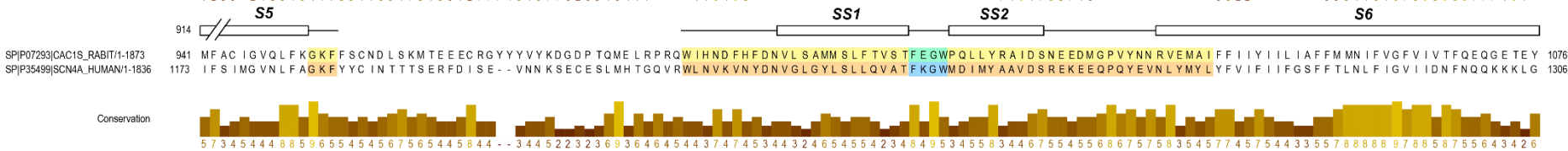
domain I



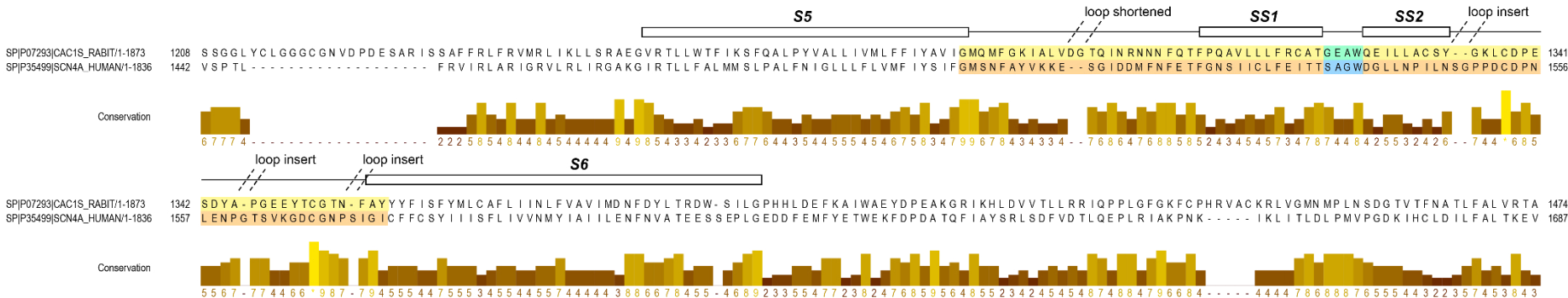
domain II



domain III

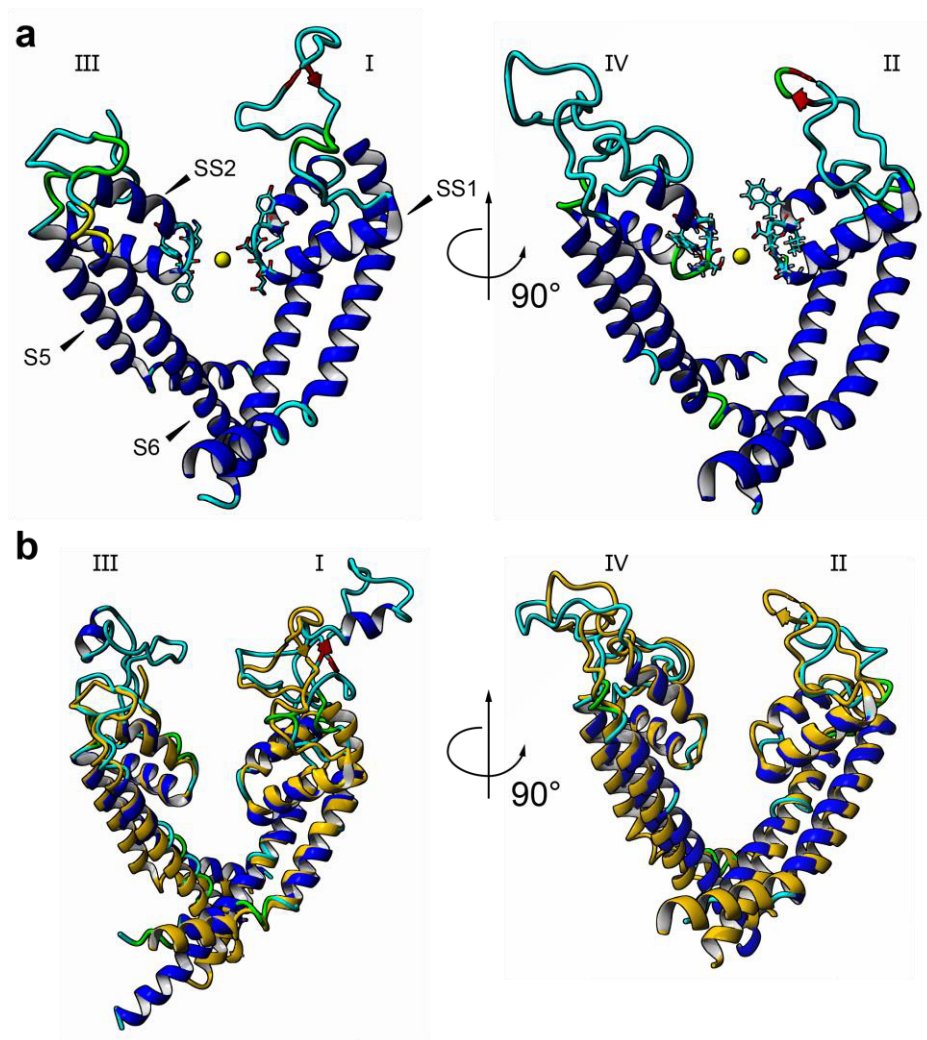


domain IV

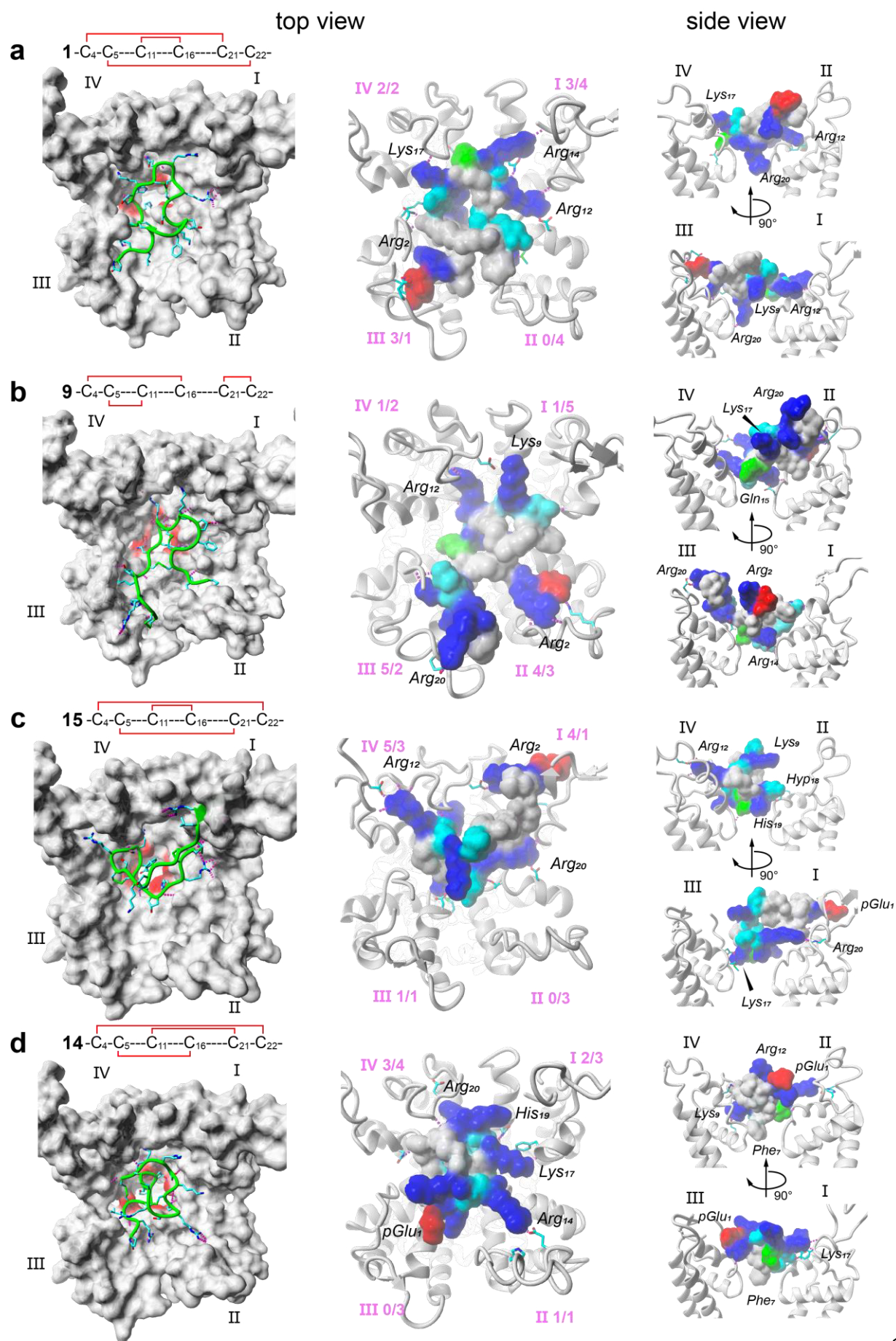


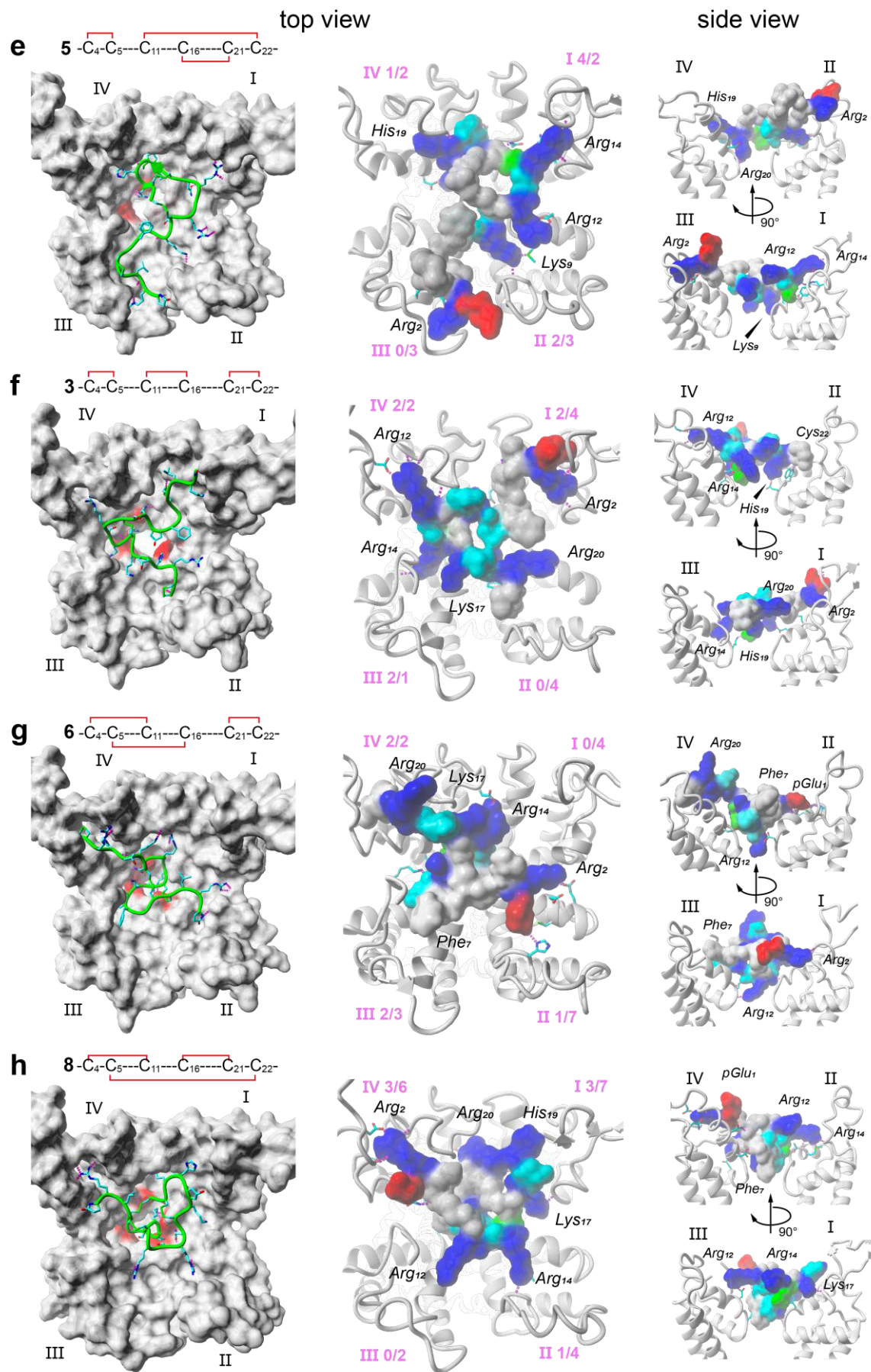
Supporting Figure 8. Alignment of the pore modules of the channel template Ca_v1.1 (PDB ID 3jbr¹⁶, uniprot ID P07293) and the target sequence of Na_v1.4 (uniprot ID P35499) of all four domains. Conservation of the aligned residues is indicated in the histogram below the sequence and given as a numerical index reflecting the conservation of physico-chemical properties in the alignment employing the AMAS method¹⁷: identities score highest and the next most conserved group contain substitutions to amino acids lying in the same physico-chemical class.

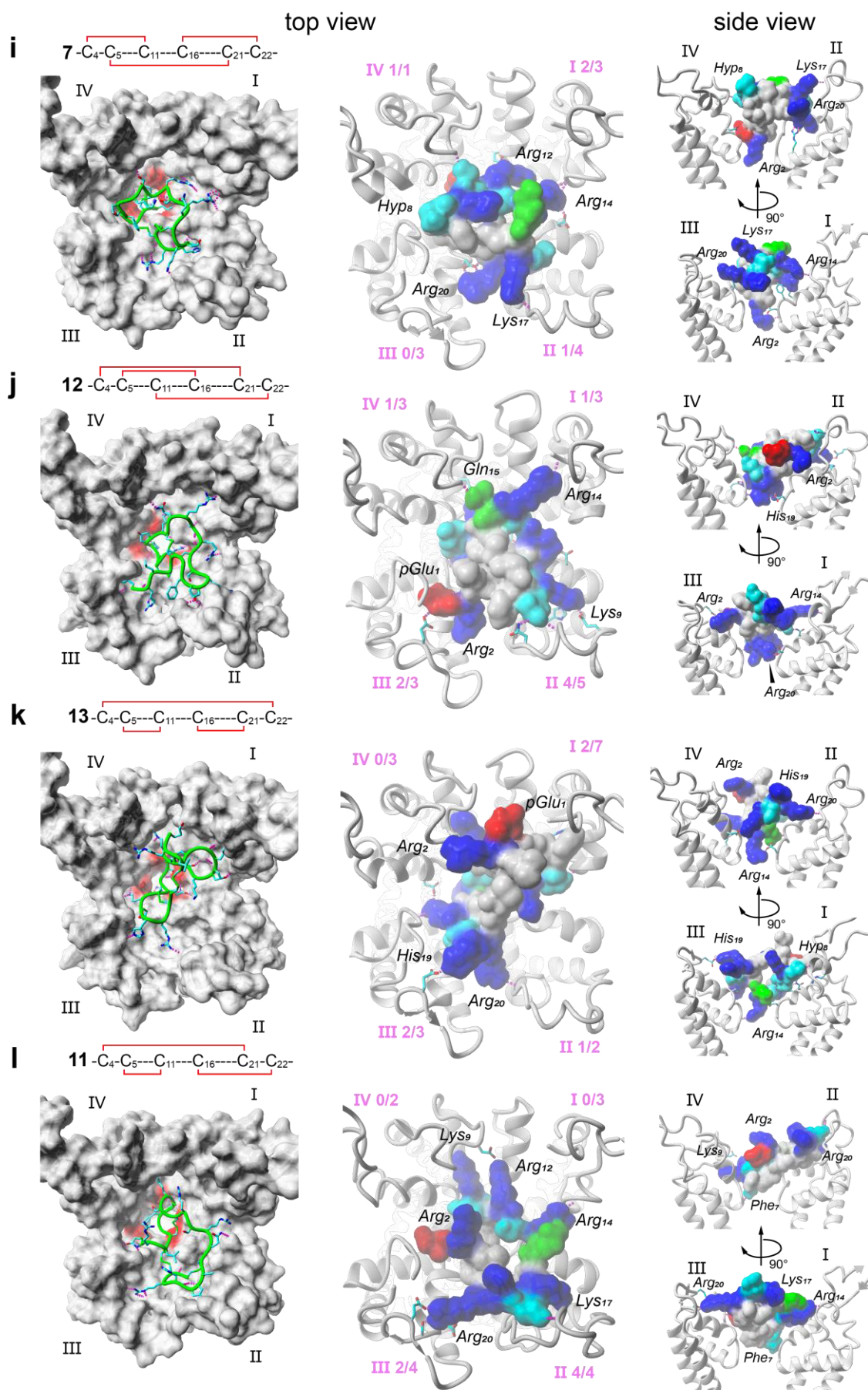
Regions from the template structure that were part of the final homology model are indicated by secondary structure elements, schematically outlined above (transmembrane segments – black box, loops - straight black line) the aligned amino acid sequences. Homology-modeled regions within the template structures are color-coded. Sequence alignments were created with Clustal Omega^{18–20} on www.uniprot.org. The alignment figure was created with JALVIEW 2.8²¹.



Supporting Figure 9. a) Homology model of Na_v1.4, selectivity filter residues are shown in stick representation (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, calcium – yellow), subunits are indicated with roman letters. b) superimposed structures of the Na_v1.4 homology model (yellow) with the recently published structure of the eukaryotic Na_vPaS channel (blue, cyan; pdb ID 5X0M)¹⁵ which show a structural deviation (r.m.s.d.) of 1.9 Å.







Supporting Figure 10. a – l) Visualization of the best-scoring μ -PIIIA-Nav1.4 complex conformations obtained from docking experiments. The four Nav1.4 domains are indicated. Left panel – topview of the toxin-channel complex. The Nav1.4 channel surface (Van der Waals) is given in gray, the selectivity filter motif (DEKA) is highlighted in red. The toxin is shown in green with side chain atoms present (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, sulfur – green). Middle and right panel – top and side view of the toxin-channel complex. The toxin is shown in surface representation (Van der Waals) and colored according to amino acid properties (coloring scheme: basic – blue, acidic – red, polar - cyan, non-polar – gray). The number of hydrogen bonds formed by the toxin with each channel domain (loops/SS1-SS2) is highlighted in purple distinguishing between hydrogen bonds formed with the TM bridging loops (S5-SS1, SS2-S6) and the pore helices (SS1-SS2).

References

- 1 S. Peigneur, M. Paolini-Bertrand, H. Gaertner, D. Biass, A. Violette, R. Stöcklin, P. Favreau, J. Tytgat and O. Hartley, *J. Biol. Chem.*, 2014, **289**, 35341–35350.
- 2 M. Góngora-Benítez, J. Tulla-Puche and F. Albericio, *Chem. Rev.*, 2014, **114**, 901–926.
- 3 I. Szabo, G. Schlosser, F. Hudecz and G. Mezo, *Biopolymers*, 2007, **88**, 20–28.
- 4 M. Muttenthaler, Y. G. Ramos, D. Feytens, A. D. De Araujo and P. F. Alewood, *Pept. Sci.*, 2010, **94**, 423–432.
- 5 K. Akaji, K. Fujino, T. Tatsumi and Y. Kiso, *Tetrahedron Lett.*, 1992, **33**, 1073–1076.
- 6 A. Cuthbertson and B. Indrevoll, *Org. Lett.*, 2003, **5**, 2955–2957.
- 7 A. A. Tietze, D. Tietze, O. Ohlenschläger, E. Leipold, F. Ullrich, T. Köhl, A. Mischo, G. Buntkowsky, M. Görlach, S. H. Heinemann and D. Imhof, *Angew. Chem. Int. Ed.*, 2012, **51**, 4058–4061.
- 8 J. Gehrmann, P. F. Alewood and D. J. Craik, *J. Mol. Biol.*, 1998, **278**, 401–415.
- 9 B. B. Carstens, G. Berecki, J. T. Daniel, H. S. Lee, K. A. V Jackson, H. S. Tae, M. Sadeghi, J. Castro, T. O'Donnell, A. Deiteren, S. M. Brierley, D. J. Craik, D. J. Adams and R. J. Clark, *Angew. Chem. Int. Ed.*, 2016, **55**, 4692–4696.
- 10 L. Poppe, J. O. Hui, J. Ligutti, J. K. Murray and P. D. Schnier, *Anal. Chem.*, 2012, **84**, 262–266.
- 11 H. Chen, D. Gordon and S. H. Heinemann, *Pflugers Arch.*, 2000, **439**, 423–432.
- 12 X. Zhang, W. Ren, P. Decaen, C. Yan, X. Tao, L. Tang, J. Wang, K. Hasegawa, T. Kumasaka, J. He, J. Wang, D. E. Clapham and N. Yan, *Nature*, 2012, **486**, 130–134.
- 13 J. Payandeh, T. Scheuer, Z. N and W. A. Catterall, *Nature*, 2011, **475**, 353–358.
- 14 J. Payandeh, T. M. G. El-Din, T. Scheuer, N. Zheng and W. A. Catterall, *Nature*, 2012, **486**, 135–139.
- 15 H. Shen, H. Shen, Q. Zhou, X. Pan, Z. Li, J. Wu and N. Yan, *Science (80-.)*, 2017, **4326**, 1–18.
- 16 J. Wu, Z. Yan, Z. Li, C. Yan, S. Lu, M. Dong and N. Yan, *Nature*, 2016, **537**, 191–196.
- 17 C. D. Livingstone and G. J. Barton, *Comput. Appl. Biol. Sci.*, 1993, **9**, 745–756.
- 18 F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson and D. G. Higgins, *Mol. Syst. Biol.*, 2011, **7**, 1–6.
- 19 H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley and R. Lopez, *Nucleic Acids Res.*, 2013, **41**, 597–600.
- 20 M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern and R. Lopez, *Nucleic Acids Res.*, 2010, **38**, 695–699.
- 21 A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp and G. J. Barton, *Bioinformatics*, 2009, **25**, 1189–1191.

Danksagung

Zunächst möchte ich mich besonders bei Prof. Dr. Diana Imhof bedanken: Vielen Dank für die enge Betreuung, das entgegengebrachte Vertrauen und die Unterstützung während meiner Doktorarbeit. Auch für das anspruchsvolle, aber schöne Thema der Conotoxine, die Erfahrungen und Kenntnisse die ich im Laufe des Projektes aneignen konnte, bin ich sehr dankbar.

Herrn PD. Dr. Jochen Winter möchte ich herzlich für die Übernahme des Koreferates danken, sowie Frau Prof. Dr. Evi Kostenis und Herrn Prof. Dr. Arne Lützen für die Mitwirkung in der Prüfungskommission und die Beteiligung am Promotionsverfahren.

Des Weiteren möchte ich mich bei den wissenschaftlichen Kooperationspartnern für die jahrelange Zusammenarbeit und zahlreichen Diskussionen bedanken. Hierbei bedanke ich mich besonders bei Dr. Alesia Tietze und Dr. Daniel Tietze für die schöne Zeit an der TU Darmstadt und die NMR-Messungen, Auswertung und Anwendung der Daten in Docking Studien, Dr. Enrico Leipold und Prof. Dr. Stefan Heinemann für die Durchführung der Bioaktivitätsmessung an der Universität Jena, Marion Schneider (Universität Bonn), Dr. Franz-Josef Mayer, Dr. Detlef Suckau, Dr. Anja Resemann (alle Bruker Daltonics Bremen) und Dr. Susanne Neupert (Universität Köln) für die Hilfe bei der Durchführung und Erweiterung meiner Massenspektrometrie-Kenntnisse, Prof. Dr. Annegret Stark für die Zusammenarbeit im IL-Projekt sowie Nils Lichtenberg und JProf. Dr. Kai Lawonn für die Zusammenarbeit in Computer Visualistics.

Allen im Arbeitskreis Imhof - Toni, Miriam, Dorle, Henning, Amelie, Ming, Charlotte, Justin und Ajay - möchte ich für die schöne Zeit innerhalb und außerhalb des Labors danken. Wir hatten viel Spaß auf Konferenzen, Retreats und bei sozialen Events. Auch den zahlreichen Praktikanten danke ich für die gute Zusammenarbeit und Mithilfe im Labor. Amelie, Henning und Charlotte danke ich besonders herzlich fürs Korrekturlesen meiner Arbeit. Ich möchte auch den gegenwärtigen und ehemaligen Kollegen im Praktikum „4. Semester Instrumentelle Analytik“ und „3. Semester Organische Chemie“ für die gute Zusammenarbeit danken.

Bei meinen Eltern und Geschwistern möchte ich mich sehr für eure Liebe und Unterstützung in meinem ganzen Leben bedanken. Liebste Carolin, Dir gebührt mein besonderer Dank für dein Verständnis und die Aufmunterung während der ganzen Zeit.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass die vorgelegte Arbeit, abgesehen von den ausdrücklich bezeichneten Hilfsmitteln, persönlich, selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde. Daten und Konzepte, die aus anderen Quellen direkt oder indirekt übernommen wurden, sind unter Angaben von Quellen kenntlich gemacht. Diese Arbeit hat in dieser oder ähnlichen Form keiner anderen Prüfungsbehörde vorgelegen und ich habe keine früheren Promotionsversuche unternommen. Für die Erstellung der vorliegenden Arbeit wurde keine fremde Hilfe, insbesondere keine entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen.

Pascal Heimer

Bonn, den 06.03.2017